(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
11 December 2003 (11.12.2003)

PCT

(10) International Publication Number
WO 03/102172 A1

(51) International Patent Classification[7]: C12N 9/00, 15/52, 15/11, C07H 21/04

(21) International Application Number: PCT/US03/17503

(22) International Filing Date: 2 June 2003 (02.06.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/385,070 31 May 2002 (31.05.2002) US

(71) Applicant *(for all designated States except US)*: **RIGEL PHARMACEUTICALS, INC.** [US/US]; 1180 Veterans Blvd., South San Francisco, CA 94080 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **ANDERSON, David** [US/US]; 2620 Tandy Turn, Eugene, OR 97401 (US). **GURURAJA, Tarikere** [IN/US]; 642 Blythe Court # 25, Sunnyvale, CA 94085 (US).

(74) Agents: **KELLY, Beth, L.** et al.; Townsend and Townsend and Crew LLP, Two Embarcadero Center, 8th Floor, San Francisco, CA 94111 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

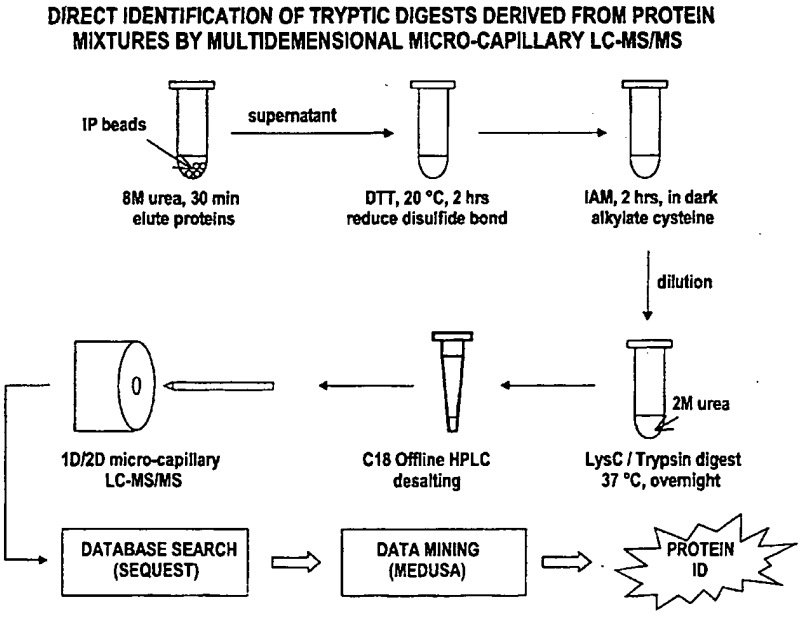
(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declaration under Rule 4.17:
— *of inventorship (Rule 4.17(iv)) for US only*

Published:
— *with international search report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: METHODS AND COMPOSITIONS FOR A UBIQUITIN MAP



(57) Abstract: The present invention provides methods for identifying protein substrates of the ubiquitin proteolytic pathway using mass spectrometry.

WO 03/102172 A1

METHODS AND COMPOSITIONS FOR A UBIQUITIN MAP

CROSS-REFERENCES TO RELATED APPLICATIONS

[0001] This application claims the benefit of priority of U.S. provisional application 60/385,070 filed May 31, 2002, which is herein incorporated by reference for all purposes. The application discloses assays for determining ubiquitination activity. Other ubiquitination assays are set forth in U.S. Application Serial Numbers 09/542,497, filed April 3, 2000; 09/826,312, filed April 3, 2001; 10/091,174, filed March 4, 2002; 10/091,139, filed March 4, 2002; 10/109,460, filed March 26, 2002; 10/108,767, filed March 26, 2002; all of which are expressly incorporated herein by reference.

BACKGROUND OF THE INVENTION

[0002] Ubiquitin is a highly conserved 76 amino acid protein expressed in all eukaryotic cells. The levels of many intracellular proteins are regulated by a ubiquitin-dependent proteolytic process. This process involves the covalent ligation of ubiquitin to a target protein, resulting in a poly-ubiquitinated target protein which is rapidly detected and degraded by the 26S proteasome.

[0003] The ubiquitination of these proteins is mediated by a cascade of enzymatic activity. Ubiquitin is first activated in an ATP-dependent manner by a ubiquitin activating enzyme (E1). The C-terminus of a ubiquitin forms a high energy thiolester bond with E1. The ubiquitin is then passed to a ubiquitin conjugating enzyme (E2; also called ubiquitin carrier protein), also linked to this second enzyme via a thiolester bond. The ubiquitin is finally linked to its target protein to form a terminal isopeptide bond under the guidance of a ubiquitin ligase (E3). In this process, chains of ubiquitin are formed on the target protein, each covalently ligated to the next through the activity of E3.

[0004] The components of the ubiquitin ligation cascade have received considerable attention (for a review, *see* Weissman, *Nature Reviews* 2:169-178 (2001)). E1 and E2 are structurally related and well characterized enzymes. There are several species of E2 (at least 25 in mammals), some of which act in preferred pairs with specific E3 enzymes to confer specificity for different target proteins. While the nomenclature for E2 is not standardized across species, investigators in the field have addressed this issue and the skilled artisan can

readily identify various E2 proteins, as well as species homologues (*See* Haas and Siepmann, *FASEB J.* 11:1257-1268 (1997)).

[0005] E3 enzymes contain two separate activities: a ubiquitin ligase activity to conjugate ubiquitin to substrates and form polyubiquitin chains via isopeptide bonds, and a targeting activity to physically bring the ligase and substrate together. Substrate specificity of different E3 enzymes is the major determinant in the selectivity of the ubiquitin-dependent protein degradation process.

[0006] Some E3 ubiquitin ligases are known to have a single subunit responsible for the ligase activity. Such E3 ligases that have been characterized include the HECT (homologous to E6-AP carboxy terminus) domain proteins, represented by the mammalian E6AP-E6 complex which functions as a ubiquitin ligase for the tumor suppressor p53 and which is activated by papillomavirus in cervical cancer (Huang et al., *Science* 286:1321-26 (1999)). Single subunit ubiquitin ligases having a RING domain include Mdm2, which has also been shown to act as a ubiquitin ligase for p53, as well as Mdm2 itself. Other RING domain, single subunit E3 ligases include: TRAF6, involved in IKK activation; Cbl, which targets insulin and EGF; Sina/Siah, which targets DCC; Itchy, which is involved in haematopoesis (B, T and mast cells); and IAP, involved with inhibitors of apoptosis.

[0007] The best characterized E3 ligase is the APC (anaphase promoting complex), which is a multi-subunit complex that is required for both entry into anaphase as well as exit from mitosis (*see* King et al., *Science* 274:1652-59 (1996) for review). The APC plays a crucial role in regulating the passage of cells through anaphase by promoting ubiquitin-dependent proteolysis of many proteins. In addition to degrading the mitotic B-type cyclin for inactivation of CDC2 kinase activity, the APC is also required for degradation of other proteins for sister chromatid separation and spindle disassambly. Most proteins known to be degraded by the APC contain a conserved nine amino acid motif known as the "destruction box" that targets them for ubiquitination and subsequent degradation. However, proteins that are degraded during G1, including G1 cyclins, CDK inhibitors, transcription factors and signaling intermediates, do not contain this conserved amino acid motif. Instead substrate phosphorylation appears to play an important role in targeting their interaction with an E3 ligase for ubiquitination (see Hershko et al., *Ann. Rev. Biochem.* 67:429-75 (1998)).

[0008] In eukaryotes, a family of complexes with E3 ligase activity play an important role in regulating G1 progression. These complexes, called SCF's, consist of at least three

subunits, SKP1, Cullins (having at least seven family members) and an F-box protein (of which hundreds of species are known) which bind directly to and recruit the substrate to the E3 complex. The combinatorial interactions between the SCF's and a recently discovered family of RING finger proteins, the ROC/APC 11 proteins, have been shown to be the key elements conferring ligase activity to E3 protein complexes. Particular ROC/Cullin combinations can regulate specific cellular pathways, as exemplified by the function of APC11-APC2, involved in the proteolytic control of sister chromatid separation and exit from telophase into G1 in mitosis (*see* King et al., *supra*; Koepp et al., *Cell* 97:431-34 (1999)), and ROC1-Cullin 1, involved in the proteolytic degradation of $I_\kappa B_\alpha$ in NF-$_\kappa$B/I_κB mediated transcription regulation (Tan et al., *Mol. Cell* 3(4):527-533 (1999); Laney et al., *Cell* 97:427-30 (1999)).

[0009] Due to the importance of ubiquitination in cellular regulation and the wide array of different possible components in ubiquitin-dependent proteolysis, there is a need to identify targets of ubiquitin-dependent proteolysis, *i.e.* ubiquitinated proteins. The present invention solves this and other problems.

BRIEF SUMMARY OF THE INVENTION

[0010] The present invention provides methods of identifying target or substrate proteins of ubiquitination pathways. Target or substrate proteins of ubiquitination pathways are ubiquitinated proteins. The present invention provides methods of generating ubiquitin maps, wherein target proteins are identified.

[0011] The present invention also provide methods of identifying a disorder associated with ubiquitin or ubiquitination. A disorder associated with ubiquitin or ubiquitination is a disease state where ubiquitination of substrate proteins is different from that of an undiseased control. The control can be from undiseased or normal tissue from the same individual, or from a different undiseased or normal individual. The differences can be in the identity of the substrates, *e.g.*, proteins that are ubiquitinated in the disease state, but not in the control; or proteins that are not ubiquitinated in the disease state, but are ubiquitinated in the control. The differences can also be in the amount of ubiquitinated protein present in the disease state. The differences can be a result of increases or decreases in ubiquitin ligation cascade activity or a result of increases or decreases in deubiquitination activity. Those of skill will recognize that ubiquitin maps are useful to determine whether a particular disease or disorder is

associated with ubiquitin. Also ubiquitin maps can be used for diagnosis of a disorder associated with ubiquitin in an individual.

[0012] In one embodiment, the present invention provides a method for the identification of ubiquitinated proteins where ubiquitinated proteins are separated from non-ubiquitinated proteins; the ubiquitinated proteins are then isolated from each other; and then the ubiquitinated proteins are identified, usually by determining the amino acid sequence of the ubiquitinated protein or peptides derived from the ubiquitinated protein and comparing them to amino acid sequence data bases. This method can be used to generate a ubiquitin map, *e.g.*, a list or compilation of ubiquitinated proteins under a given condition or in a particular cell type or subject.

[0013] In one aspect ubiquinated proteins are identified in a population of cells and compared to ubiquitinated proteins in a second population of cells. The cell populations can come from different disease states, *e.g.*, cancer cells vs normal cells; or one cell population can be treated with a modulator of a ubiquitination pathway and compared to an untreated cell population.

[0014] In one aspect, the isolation of ubiquitinated proteins from each is done using 2D gel analysis. In another aspect the isolation of ubiquitinated proteins from each other is done by purifying the proteins.

[0015] In one aspect, the ubiquitinated proteins are labeled with a label that is selective for ubiquitin. The ubiquitinated proteins are then separated from non-ubiquitinated proteins by purification using a complementary binding partner that binds to the labeled ubiquitin. The ubiquitin labeling can occur *in vivo* or *in vitro*.

[0016] For *in vitro* labeling, cell extracts are first prepared and then labeled ubiquitin is added to the extracts and ubiquitination of target or substrate proteins in the cell extracts is allowed to occur, thus generating labeled ubiquitinated target or substrate proteins. Labels include *e.g.*, a His-tag, Glutathione-S-transferase, Flag-tag, HA-tag, affinity tag, epitope tag and biotin. The *in vitro* reactions can be performed in the presence of proteosome inhibitors, for example MG132, PS341 and epoxymicin. The *in vitro* reactions can be performed in the presence of deubiquitinase inhibitors, for example ubiquitin aldehyde. The *in vitro* reactions can be performed using E1 ubiquitin activating enzyme and E2 ubiquitin conjugating inhibitors, for example iodoacetamide. The *in vitro* reactions can also include an ATP regenerating system.

[0017] For *in vivo* labeling, cells are provided with a labeled ubiquitin moiety and ubiquitin targets are labeled *in vivo* with the labeled ubiquitin moiety. In some aspects the labeled ubiquitin is a recombinant protein, that is encoded by a nucleic acid that has been transduced into the cells. Labels include *e.g.*, a His-tag ubiquitin, Glutathione-S-transferase ubiquitin and Flag-tag ubiquitin HA-tagged ubiquitin, affinity tag ubiquitin, epitope tag ubiquitin and biotin tagged ubiquitin. The *in vivo* reactions can also be performed by incubating the cells with proteosome inhibitors, for example MG132, PS341 and epoxymicin. The *in vivo* reactions can also be performed by incubating the cells with deubiquitinase inhibitors, for example ubiquitin aldehyde. The *in vivo* reactions can be performed using E1 ubiquitin activating enzyme and E2 ubiquitin conjugating inhibitors, for example iodoacetamide.

[0018] In a further embodiment, ubiquitinated proteins are separated by affinity purification, using ubiquitin or a label on the ubiquitin for the affinity purification step, followed by a separation in a second dimension using a combination of chromatography and mass spectroscopy, *e.g.*, LC-MS/MS. In one aspect, the affinity purified ubiquitinated proteins are cleaved before the LC-MS/MS step is performed. The cleavage can be done using a protease, for example lys C endo protease or trypsin.

[0019] The present invention also provides methods for diagnosing a disorder associated with ubiquitination by obtaining or generating a ubiquitin map from cells from an individual suspected of having a disorder associated with ubiquitin. The ubiquitin map is compared to a second ubiquitin map from a control where a difference in identified ubiquitinated protein(s) is indicative of a disorder associated with ubiquitination.

[0020] The present invention provides a method of identifying a disorder associated with ubiquitin, by obtaining or generating a ubiquitin map from cells from an individual with a disorder suspected to be associated with ubiquitin and comparing that ubiquitin map to a second ubiquitin map from a control, whereby a difference in identified ubiquitinated protein(s) is indicative of a disorder associated with ubiquitination. Ubiquitin maps can be generated from a plurality of individuals with the disorder suspected to be associated with ubiquitin.

[0021] The present invention also provides a method of identifying a target protein of a ubiquitin pathway protein, using a modulator of the ubiquitin pathway protein. As before, ubiquitinated proteins are first separated from non-ubiquitinated proteins; isolated from each

other; and then identified. A difference in ubiquitination in the absence of the modulator indicates thata protein is a target for a ubiquitin pathway protein.

[0022] The ubiquitin pathway protein can be ubiquitin ligation cascade component, *e.g.*, an E1, E2, or E3 protein. The modulator can be an inhibitor of the ubiquitin ligation cascade component, *e.g.*, a dominant negative mutant, small molecules, siRNA, and antisense RNA. The modulator can be an activator of the ubiquitin ligation cascade component, *e.g.*, an overexpressed ubiquitin ligation cascade component.

[0023] The ubiquitin pathway protein can be a deubiquitinating protein. The modulator can be an inhibitor of the deubiquitinating protein, *e.g.*, a dominant negative mutant, small molecules, siRNA, and antisense RNA. The modulator can be an activator of the deubiquitinating protein, *e.g.*, an overexpressed deubiquitinating protein.

BRIEF DESCRIPTION OF THE DRAWINGS

[0024] Figure 1 depicts elements of ubiquitin-mediated proteolysis.

[0025] Figure 2 depicts *in vitro* ubiquitination of cell lysate proteins using his6-ubiquitin as a label, visualized as an anti-his tag Western blot of cell lysate incubated with different combinations of reagents. HeLa cells were lysed in 0.2% NP-40 and incubated with N-his6-tagged ubiquitin, and when present, an ATP regenerating system, the ubiquitin protease inhibitor ubiquitin aldehyde, the proteasome inhibitor MG132, and the E1 ubiquitin activating enzyme and E2 ubiquitin conjugating inhibitor iodoacetamide. Lanes 1-4, 30 min. incubation of reagents in cell lysate followed by SDS-PAGE and Western blotting. Lane 1, incubation with his-ubiquitin, ubiquitin aldehyde, MG132, but no ATP recycling system. Few higher molecular weight his-ubiquitin bands are detected. The major band at ca. 9 kDa is his-ubiquitin. Lane 2, incubation with these reagents, the ATP recycling system, and the E1 and E2 inhibitor iodoacetamide. This is the full in vitro ubiquitin labeling system, but with the E1 and E2 inhibitor iodoacetamide added. This inhibitor abolishes the generation of higher molecular weight his-ubiquitin bands. Lane 3, incubation as in 2 without iodoacetamide; this is the complete *in vitro* ubiquitin labeling system. Lane 4, incubation with all reagents in 3 except ubiquitin aldehyde and MG132. Lanes 1'-4'. These lanes show the 120 min. time point for the same reactions as in 1-4. The intensity of bands in lane 3' has increased but the other lanes still have few bands; the creation of higher molecular weight bands is time dependent.

[0026] Figure 3 depicts detection of ubiquitinated proteins in Hela cell lysates and nickel-chelate affinity chromatography extracts of in vitro labeled cells. Left panels, blot of cell lysates 120 min. after addition of his6-ubiquitin and the ATP regeneration system. Anti-his tag and anti-ubiquitin antibodies were used in the absence (lane 1) and presence (lane 2) of ATP and the ATP regenerating system. Ladders of proteins cross-reactive with both the anti-his tag and anti-ubiquitin antibodies are visible. Right panels, affinity extracts using nickel chelate chromatography. A silver stained gel of in vitro labeled cellular proteins shows extraction of more proteins in the presence than in the absence of ATP and the ATP regeneration system. Anti-his tag and anti-ubiquitin Western blots show the same result.

[0027] Figure 4 depicts direct identification of tryptic digest derived from protein mixtures by multidimensional micro-capillary LC-MS/MS.

[0028] Figure 5 depicts data processing and bioinformatics used to identify ubiquitinated target proteins.

DETAILED DESCRIPTION OF THE INVENTION

INTRODUCTION

[0029] The ubiquitin ligase cascade functions to regulate protein turnover in a cell by closely regulating the degradation of specific proteins. By regulating protein degradation, cells can quickly eliminate a protein that in turn regulates another function (*e.g.*, a transcription factor that is needed to express a particular gene). Furthermore, this form of control is very effective, as the elimination of a particular protein ensures that the process governed by the protein is shut-down. In this process, ubiquitin or a ubiquitin-like protein (see Figure 4) is covalently ligated to a target or substrate protein, resulting in a poly-ubiquitinated target protein that is rapidly detected and degraded by the 26S proteasome.

[0030] The timed ubiquitination and subsequent destruction of proteins have been implicated in various cellular processes including cell cycle (activation of cyclinE/Cdk2); inflammation (NFkB activation); angiogenesis (oxygen sensing and VEGF production) and cancer (the action of anti-mitotics at anaphase). (*See, e.g.*, Shah *et al., Surg. Oncol.* 10:43-52 (2001)). The hallmark of this ubiquitin-mediated proteolytic pathway is the covalent attachment of the 76-residue ubiquitin (Ub) polypeptide to lysine residues of substrate proteins *via* an E1/E2/E3 enzyme cascade (E1: Ub-activating enzyme, E2: Ub conjugating

enzyme, and E3: Ub ligase). (*See, e.g.*, Hershko *et al.*, *Annu. Rev. Biochem.* 67:425-479 (1998)). The ubiquitination of proteins.is also affected by a deubiquitination pathway.

[0031] A number of disease states are influenced by ubiquitination of specific target proteins, *e.g.*, inflammation and inflammatory diseases, immune system diseases, infectious diseases, including bacterial and viral diseases, and cancer and other proliferative diseases. As an example, p53 is an important tumor suppressor. In cells, p53 functions as a DNA-binding transcription factor which induces the expression of genes involved in DNA repair, apoptosis, and the arrest of cell growth. In approximately 50 % of all human cancer p53 is. inactivate by deletion or mutation. The level of p53 in the cell is maintained at low steady-state levels, and is induced and activated post-translationally by various signal pathways responsive to cellular stress (Lakin *et al.* (1999) *Oncogene* 18:7644-7655; Oren, M. (1999) *J. Biol. Chem* 274:36031-36,034). Stimuli that trigger the stress response and activate p53 include oxygen stress, inappropriate activation of oncogenes and agents that cause damage to DNA (*e.g.*, ionizing radiation, chemicals, and ultra violet light). Mdm2 belongs to the second subclass of single subunit E3 ligating agents and is involved in regulating the function and stability of p53. p53 levels are thus also regulated by uiquitination and those of skill will recognize that the ubiquitination pathway that regulates p53 is a target for drugs that modulate p53 activity.

[0032] The present invention provides a method to identify proteins that are targeted by ubiquitin pathways. The present invention also provides methods to determine whether a disease or disorder is associated with ubiquitination. The present invention also provides methods to diagnose an individual as having a disease or disorder associated with ubiquitin. Identification of targets of the ubiquitin-mediated proteolysis will allow analysis of disease states that are influenced by the activity of ubiquitin pathways and of drugs used to treat such disease states. For example, a map of ubiquitinated cellular proteins is useful for determining the specific protein affected by inhibitors of E1, E2 or E3 enzymes. The present invention provides strategies to isolate polyubiquitinated substrate proteins, for identification by mass spectrometry followed by amino acid sequence analysis of the ubiquitinated proteins.

DEFINITIONS

[0033] "Ubiquitin ligation pathway, cascade, or component" refers to ubiquitin and ubiquitin-like molecules, and E1, E2, and E3 proteins and their substrates, which are involved in the ubiquitination process (*see, e.g.*, Weissman, *Nature Reviews* 2:169-178 (2001); *see*

also WO 01/75145)). A "deubiquitinating agent, protein or polypeptide" refers to a protein that catalyzes the binding, release, or cleavage of a ubiquitin moiety from a ubiquitin complex. A "ubiquitination pathway," as used herein, encompasses both a ubiquitin ligation pathway, cascade, or component and a deubiquitinating agent, protein or polypeptide.

[0034] A "target or substrate protein" of a ubiquitination pathway; a ubiquitin ligation pathway, cascade, or component; or a deubiquitinating agent, protein or polypeptide refers to a ubiquitinated protein, polypeptide or a ubiquitin complex.

[0035] By "ubiquitin" herein is meant a polypeptide which is ligated to another polypeptide by ubiquitin ligase enzymes. The ubiquitin can be from any species of organism, preferably a eukaryotic species. Preferably, the ubiquitin is mammalian. More preferably, the ubiquitin is human ubiquitin. In a most preferred embodiment, the ubiquitin has the amino acid sequence depicted in figure 15A.

[0036] In a preferred embodiment, when ubiquitin is ligated to a target protein, that protein is targeted for degradation by the 26S proteasome.

[0037] Preferred embodiments of the invention utilize a 76 amino acid human ubiquitin. Other embodiments utilize variants of ubiquitin, as further described below.

[0038] Also encompassed by "ubiquitin" is naturally occurring alleles and man-made variants of such a 76 amino acid polypeptide. In a preferred embodiment, the ubiquitin has the amino acid sequence of that depicted in ATCC accession number P02248, incorporated herein by reference. ATCC accession numbers are found in Genbank. Sequences of GenBank accession numbers are incorporated herein by reference. GenBank is known in the art, see, *e.g.*, Benson, DA, *et al., Nucleic Acids Research* 26:1-7 (1998) and www.ncbi.nlm.nih.gov/.

[0039] In a preferred embodiment, ubiquitin has a tag, as defined herein, with the complex being referred to herein as "tag-Ub". Preferred ubiquitin tags include, but are not limited to, labels, partners of binding pairs and substrate binding elements. In a most preferred embodiment, the tag is a His-tag or GST tag.

[0040] The present invention provides methods and compositions comprising combining ubiquitin and E1. By "E1" is meant a ubiquitin activating enzyme. In a preferred embodiment, E1 is capable of transferring ubiquitin to an E2, defined below. In a preferred

embodiment, E1 binds ubiquitin. In a preferred embodiment, E1 forms a high energy thiolester bond with ubiquitin, thereby "activating" the ubiquitin.

[0041] In some embodiments a ubiquitin map is generated from cells that have increased expression of an E1 protein, that have decreased expression of an E1 protein, that express a dominant negative mutant of an E1 protein, or that have been treated with a modulator of E1 protein activity. In a preferred embodiment, E1 proteins useful in the invention include those having the amino acid sequence of the polypeptide having ATCC accession numbers A38564, S23770, AAA61246, P22314, CAA40296 and BAA33144, incorporated herein by reference.

[0042] In a preferred embodiment, nucleic acids which may be used for producing E1 proteins for the invention include, but are not limited to, those disclosed by ATCC accession numbers M58028, X56976 and AB012190, incorporated herein by reference. Variants of the cited E1 proteins, also included in the term "E1", can be made as described herein.

[0043] In a preferred embodiment, the compositions of the invention comprise E2. By "E2" is meant a ubiquitin carrier enzyme (also known as a ubiquitin conjugating enzyme). In a preferred embodiment, ubiquitin is transferred from E1 to E2. In a preferred embodiment, the transfer results in a thiolester bond formed between E2 and ubiquitin. In a preferred embodiment, E2 is capable of transferring ubiquitin to an E3, defined below. In a preferred embodiment, the ubiquitination substrate protein is ubiquitin.

[0044] In some embodiments a ubiquitin map is generated from cells that have increased expression of an E2 protein, that have decreased expression of an E2 protein, that express a dominant negative mutant of an E2 protein, or that have been treated with a modulator of E2 protein activity. In a preferred embodiment, proteins which may be used in the present invention as E2 include, but are not limited to, those having the amino acid sequences disclosed in ATCC accession numbers AAC37534, P49427, CAA82525, AAA58466, AAC41750, P51669, AAA91460, AAA91461, CAA63538, AAC50633, P27924, AAB36017, Q16763, AAB86433, AAC26141, CAA04156, BAA11675, Q16781, NP_003333, BAB18652, AAH00468, CAC16955, CAB76865, CAB76864, NP_05536, 000762, XP_009804, XP_009488, XP_006823, XP_006343, XP_005934, XP_002869, XP_003400XP_009365, XP_010361, XP_004699, XP_004019, O14933, P27924, P50550, P52485, P51668, P51669, P49459, P37286, P23567, P56554, and CAB45853, each of which is incorporated herein by reference. Other preferred E2 proteins are described in USSN

10/152,156, filed May 20, 2002, and incorporated by reference herein in its entirety. Particularly preferred are sequences disclosed in ATCC accession numbers NP003331, NP003330, NP003329, P49427, AAB53362, NP008950, XP009488 and AAC41750, also incorporated by reference. The skilled artisan will appreciate that many different E2 proteins and isozymes are known in the field and may be used in the present invention, provided that the E2 has ubiquitin conjugating activity. Also specifically included within the term "E2" are variants of E2, which can be made as described herein.

[0045] In a preferred embodiment, nucleic acids which may be used to make E2 include, but are not limited to, those nucleic acids having sequences disclosed in ATCC accession numbers L2205, Z29328, M92670, L40146, U39317, U39318, X92962, U58522, S81003, AF031141, AF075599, AJ000519, XM009488, NM007019, U73379, L40146 and D83004, each of which is incorporated herein by reference. As described above, variants of these and other E2 encoding nucleic acids may also be used to make variant E2 proteins.

[0046] The present invention provides methods and compositions comprising E3. By "E3" is meant a ubiquitin ligase, comprising one or more components associated with ligation of ubiquitin to a ubiquitination substrate protein for ubiquitin-dependent proteolysis.

[0047] In some embodiments a ubiquitin map is generated from cells that have increased expression of an E3 protein, that have decreased expression of an E3 protein, that express a dominant negative mutant of an E3 protein, or that have been treated with a modulator of E3 protein activity.

[0048] In a preferred embodiment, RING finger proteins include, but are not limited to, proteins having the amino acid sequence disclosed in ATCC accession numbers AAD30147 and AAD30146 and 6320196, incorporated herein by reference. Other preferred RING finger proteins are described in USSN 10/152,156, filed May 20, 2002, and incorporated by reference herein in its entirety.

[0049] In a preferred embodiment, Cullins include, but are not limited to, proteins having the amino acid sequences disclosed in ATCC accession numbers 4503161, AAC50544, AAC36681, 4503163, AAC51190, AAD23581, 4503165, AAC36304, AAC36682, AAD45191, AAC50548, Q13620, 4503167 and AAF05751, each of which is incorporated herein by reference. In addition, in the context of the invention, each of the RING finger proteins and Cullins encompass variants of the known or listed sequences, as described herein.

[0050] These E3 proteins and variants may be made as described herein. In a preferred embodiment, nucleic acids used to make the RING finger proteins include, but are not limited to, those having the nucleic acid sequences disclosed in ATCC accession numbers AF142059, AF142060 and nucleic acids 433493 to 433990 of NC 001136. In a preferred embodiment, Cullins are made from nucleic acids including, but not limited to, those having nucleic acid sequences disclosed in ATCC accession numbers NM 003592, U58087, AF062536, AF126404, NM 003591, U83410, NM 003590, AB014517, AF062537, AF064087, AF077188, U58091, NM 003478, X81882 and AF191337, each of which is incorporated herein by reference. As described above, variants of these sequences are also encompassed by the invention.

[0051] E3 may also comprise other components, such as SKP1 and F-box proteins. The amino acid and nucleic acid sequences for SKP1 are found in ATCC accession numbers AAC50241 and U33760, respectively. Many F-box proteins are known in the art and their amino acid and nucleic acid sequences are readily obtained by the skilled artisan from various published sources.

[0052] Ubiquitination of proteins is also affected by deubiquitination agents, *e.g.*, polynucleotides and polypeptides. Ubiquitin-mediated proteolytic processing is catalyzed by deubiquitinating agents. The deubiquitinating agents, for example DeUbiquitinating (DUB) enzymes, are cysteine proteases that can hydrolyze either the ϵ-linked isopeptide bond or α-linked peptide bond at the C-terminus of a ubiquitin. In general, the deubiquitinating agents can specifically cleave ubiquitin complexes having the structure ubiquitin-N (Ub-N), where N is any number of leaving groups ranging in size, for example, from small amines and thiols to ubiquitin moieties and other proteins. These deubiquitinating agents can process polyubiquitin chains to generate free ubiquitin from precursor fusion polypeptides; affect pools of free ubiquitin by recycling branched chain ubiquitin and, also, remove ubiquitin from polyubiquitin or monomeric ubiquitin attached to a target protein (Johnston *et al.* (1999) *EMBO* 18:3877-3887; Johnston *et a*l. (1997) *EMBO* 16:3787-3796), and specifically cleave the ubiquitin moiety of a chimeric ubiquitin fusion polypeptide (Dang *et al.* (1998) *Biochemistry* 37:1868-1879). In addition, ubiquitin-like proteases also process ubiquitin in the same manner as the cysteine proteases (Olvera and Wool (1993) *J. Biol. Chem.* 268:17967-17974); Haas *et al.* (1996) *Mol. Cell. Biol.* 35:5385-5394; Matunis *et al.* (1996) *J. Cell. Biol.* 135:1457-1470; Narasimhan *et al.* (1996) *J. Biol. Chem.* 271:324-330; Mahajan *et al.* (1997) *Cell* 88:97-107).

[0053] Ubiquitin is synthesized in various forms that are functionally distinct. For example, ubiquitin is synthesized as a linear head-to-tail polyubiquitin precursor. Release of the monomeric ubiquitin by a deubiquitinating agent involves specific enzymatic cleavage between residues of the fused ubiquitin moieties of the polyubiquitin precursor. The last ubiquitin moiety in many of these precursors is encoded with an extra C-terminal residue that can be removed to expose the active C-terminal Gly. In general, the deubiquitinating agents involved in the recycling of ubiquitin are thiol proteases that recognize the C-terminal domain or C-terminal residue of ubiquitin. The deubiquitinating agents can be divided into two classes: ubiquitin C-terminal hydrolases (UCH) (Pickart and Rose (1985) *J. Biol. Chem.* 261:10210-10217) and ubiquitin-specific proteases (UBP; isopeptidases) (Tobias and Varshavsky (1991) 266:12021-12028).

[0054] UBPs contain six conserved domains, including a domain called the "CYS box" containing a conserved cysteine, a domain containing a conserved aspartic acid, and a domain called the "HIS box" containing a conserved histidine, which distinguish members of the UBP class. In particular, the domain containing the cysteine residue and domain containing the histidine residue have short sequences flanking these residues which are highly conserved in UBPs. Some members of the UBP class contain multiple ubiquitin binding sites, for example, DUB1, isoT, UBP3, Doa4, Tre2, and FAF. In addition, some members of the UBP class are transcriptionally induced in response to cytokines. The members of the UCH class are also cysteine proteases. However, members of this class do not contain the six conserved domains characteristic of the UBP class. Further, members of the UCH class contain only one ubiquitin binding site, and preferentially cleave ubiquitin from small molecules, for example, peptides and amino acids. In addition, the two classes of deubiquitinating agents share little sequence homology.

[0055] The function of the deubiquitinating enzymes include, for example, the disassembly of polyubiquitin to recycle ubiquitin, releasing ubiquitin from 26S proteasome substrates, releasing monomeric ubiquitin from ubiquitin fusion polypeptide precursors, the reversal of regulatory ubiquitinatination (e.g. the stabilization of protein substrates), the editing of ubiquitinated proteins that have been inappropriately unbiquitinated proteins, and regenerating active ubiquitin from adducts with small nucleophiles (*e.g.*, glutathione) that may be generated by side reactions (Wilkinson and Hoschstrasser (1998) In Peters, J.M., Harris, J.R. and Finley, D. (Eds), Ubiquitin and the Biology of the Cell. Plenum Press, New York, NY, pp. 99-125). The end result of each of these activities can affect the level of free

ubiquitin and other specific proteins in the cell (D'Andrea *et al.* (1998) *Critical Reviews In Biochemistry and Molecular Biology* 33:337-352).

[0056] For example, the yeast UBP14p deubiquitinating agent and its human homolog, isopeptidase-T, hydrolyze free polyubiquitin chains and promote the degradation of polyubiquitinated protein substrates by the 26S proteasome. One of the functions of isopeptidase-T in cells is thought to be the disassembly of unanchored polyubiquitin chains and sequential degradation of the polyubiquitin chains into ubiquitin monomers.

[0057] Further the yeast Doa4 deubiquitinating agent promotes ubiquitin-mediated proteolysis of cellular substrates. In particular, Doa4 appears to function in the hydrolysis of isopeptide-linked polyubiquitin chains from peptides that are the by-products of proteasome degradation. In addition, Doa4 appears to function in the cleavage of polyubiquitin from peptide degradation products. In general, the isopeptidases can produce free monomeric ubiquitin from branched or linear polyubiquitin chains, and from ubiquitin or polyubiquitin attached to target proteins or attached to degradation products or remnants of the ubiquitin substrate, for example, peptides or amino acids.

[0058] Deubiquitinating agents that promote stabilization of substrates include the FAF protein, which deubiquitinates and rescues a ubiquitin-conjugated target protein from degradation by the proteasome. The PA700 isospeptidase, another deubiquitinating agent, also prevents proteasome degradation apparently by removing ubiquitin moieties, one at a time, beginning from the distal end of a polyubiquitin chain.

[0059] Deubiquitinating agents are also associated with growth control. For example, the mammalian oncoprotein Tre-2 is a member of the UBP class of deubiquitinating agents. The truncated UPB lacking the histidine domain and lacking deubiquitinating activity is the transforming isoform of the Tre-2 oncoprotein, while, the full-length Tre-2 protein has deubiquitinating activity but does not have transforming activity. The full-length Tre-2 protein is thought to act as an intracellular growth suppressor. DUB-1 is another UBP that is thought to regulate cellular processes. DUB-1 is induced by interleukin-3 stimulation. In general, members of this class of deubiquitinating agents are thought to be responsive to cytokines. Further, DUB-2, another member of this class, is induced by interleukin-2. (Zhu *et al.* (1997) *J. Biol. Chem.* 272:51-57). This class of deubiquitinating agents may deubiquitinate cell surface growth factor receptors, thereby, prolonging receptor half life and amplifying growth signals; and may also deubiquitinate proteins involved in signal

transduction and proteins that are cell cycle regulators, for example, cyclins and cyclin-CDK inhibitors.

[0060] UBPs are known to be involved in the chromatin regulatory process and transcriptional silencing. For example, UBP-3 forms a complex with SIR-4, a trans-acting factor that is required for activating and maintaining transcriptional silencing. Consequently, UBP-3 is thought to act as an inhibitor of transcriptional silencing by stabilizing an inhibitor or by removing a positive regulator. As a further example, the murine UNP protooncogene encodes a nuclear ubiquitin protease that when overexpressed results in oncogenic transformation in NIH3T3 cells. The cDNA corresponding to the human homolog of the murine UNP protooncogene was cloned and mapped to a chromosomal region frequently rearranged in human tumor cells. Moreover, the levels of the protooncogene were elevated in small cell tumors and adenocarcinomas of the lung. Thus, the this protooncogene may have a causative role in the neoplastic process (Gray *et al.* (1995) *Oncogene* 10:2179-2183).

[0061] Another UBP designated UBP-43, was cloned from a leukemia fusion protein in AML1-ETO Knockout mice, and has been shown to function in hematopoitic cell differentiation. The overexpression of this gene blocks cytokine-induced terminal differentiation of monocytic cells (Liu *et al.* (1999) *Molecular and Cellular Biology* 19:3029-3038).

[0062] Examples of deubquitinating agents and proteins are found in USSN 10/232,759, filed August 30, 2002, and incorporated by reference herein in its entirety.

[0063] As used herein, "deubiquitinating activity" refers to any biological activity associated with a deubiquitinating agent known in the art, for example, a cellular process, catalytic property, and more specifically, the binding, release, or cleavage of a ubiquitin moiety from a ubiquitin complex. Examples of cellular processes involving deubiquitinating agents include, but are not limited to, the disassembly of polyubiquitin to recycle ubiquitin; releasing of ubiquitin from 26S proteasome substrates, releasing of monomeric ubiquitin from ubiquitin fusion polypeptide precursors, reversal of regulatory ubiquitinatination, editing of ubiquitinated proteins that have been inappropriately unbiquitinated proteins, and regeneration of active ubiquitin from adducts with small nucleophiles (*e.g.*, glutathione) that may be generated by side reactions (Wilkinson and Hoschstrasser, 1998, incorporated herein by reference). Generally, such deubiquitinating activity affects the level of free ubiquitin and

other specific proteins in the cell (D'Andrea *et al.* (1998) *Critical Reviews In Biochemistry and Molecular Biology* 33:337-352, incorporated herein by reference).

[0064] A preferred deubiquitinating activity is the release of a ubiquitin moiety from either a polyubiquitin chain or protein substrate.

[0065] Deubiquitinating activity, or the modulation of deubiquitinating activity, can be detected and measured using the methods described herein or known in the art (*e.g., see* Sjolander *et al.* (1991) *Anal. chem.* 63:2338-2345; Szabo *et al.* (1995) *Curr. Opin. Struct. Biol.* 5:699-705; and U.S. Pat. Ser. No. 6,329,171 to Kapeller-Libermann *et al.*; Zhu *et al.* (1997) *Journal of Biological Chemistry* 272:51-57, Mitch *et al.* (1999) *American Journal of Physiology* 276: C1132-C1138; Liu *et al.* (1999) *Molecular and Cell Biology* 19:3029-3038; Ciechanover *et al.* (1994) *The FASEB Journal* 8:182-192; Chiechanover (1994) *Biol. Chem.* Hoppe-Seyler 375:565-581; Hershko *et al.* (1998) *Annual Review of Biochemistry* 67:425-479; Swartz (1999) *Annual Review of Medicine* 50:57-74, Ciechanover (1998) *EMBO Journal* 17:7151-7160; and D'Andrea *et al.* (1998) *Critical Reviews in Biochemistry; and Molecular Biology* 33:337-352). Examples of assays for the detection and measurement of deubiquitinating activity include, but are not limited to, the disappearance of ubiquitinated polypeptides (*i.e.*, ubiquitin complexes), including decrease in the amount of polyubiquitin or ubiquitinated protein or protein remnant or fragment; appearance of intermediate and end products of deubiquitinating activity, *e.g.*, the appearance of free ubiquitin monomers or released or cleaved ubiquitin moiety; general or specific protein turnover; binding to ubiquitin moiety; binding to ubiquitinated polypeptides (*i.e.*, ubiquitin complexes); interaction with ATP or cellular components such as trans-acting regulatory factors; and stabilization of specific proteins. Assays for deubiquitination activity are also found in USSN 10/232,759, filed August 30, 2002, and incorporated by reference herein in its entirety.

[0066] In an aspect of the present invention, the ubiquitin complex comprises the general structure Ub-N, where Ub is a ubiquitin moiety and attached to N via an isopeptide or peptide bond, and N can be any number of leaving groups ranging from a small amine or thiol to another ubiquitin moiety or another protein. For example, N can be a ubiquitin substrate; and Ub-N can be a cleavable ubiquitin fusion polypeptide. Examples of cleavable ubiquitin fusion polypeptides include, but is not limited to, a ubiquitin moiety fused to another ubiquitin moiety or another polypeptide; or a branched ubiquitin peptide. The ubiquitin moiety can comprise a full-length ubiquitin polypeptide or a peptide encoding a subsequence

of the full-length amino acid sequence of a ubiquitin polypeptide that can be specifically cleaved by a deubiquitinating agent. In one aspect, the ubiquitin moiety comprises a the C-terminus of a ubiquitin moiety. Examples of ubiquitin substrates include, but are not limited to, a ubiquitin agent, a target protein, or a mono- or poly-ubiquitin moiety which is preferably attached to a ubiquitin agent or target protein.

[0067] "Identification of ubiquitinated proteins" refers to the determination of the amino acid sequence of a ubiquitinated protein or the sequence of peptides derived from a ubiquitinated protein. In some embodiments, the amino acid sequences will be compared to amino acid sequence data bases, as described herein. In a preferred embodiment, the amino acid sequence of a ubiquitinated protein is determined using mass spectrometry.

[0068] "Separating ubiquitinated proteins from non-ubiquitinated proteins" refers to a process of detecting ubiquitinated proteins and then physically removing them from non-ubiquitinated proteins. The separation can be done using antibodies or proteins or protein domains that bind specifically to polyubiquitin moieties found on substrates of the ubiquitin ligation pathway. The antibodies or other proteins are preferably bound to a substrate and after binding to the ubiquitinated proteins, the non-ubiquitinated proteins can be washed or otherwise removed. The ubiquitin can also include a tag or label and antibodies to the tag or label or other binding moieties, *e.g.*, nickel NAT agars, for a polyhistidine tag, can be used to effect the separation.

[0069] "Isolating ubiquitinated proteins from each other" refers to physically separating ubiquitinated proteins from each other. A variety of methods can be used to isolate ubiquitinated proteins, including gel electrophoresis, *e.g.*, both 1 dimensional and 2 dimensional polyacrylimide gel electrophoresis, and purification techniques. In a preferred embodiment ubiquitinated proteins are isolated from each other using liquid chromatography, including HPLC and reverse phase chromatography.

[0070] A "disorder associated with ubiquitin or ubiquitination" refers to a disease state where ubiquitination of substrate proteins is different from that of an undiseased control. The control can be from a different undiseased individual or from undiseased tissue in the same individual. The differences can be in the identity of the substrates, *e.g.*, proteins that are ubiquitinated in the disease state, but not in the control; or proteins that are not ubiquitinated in the disease state, but are ubiquitinated in the control. The differences can also be in the amount of ubiquitinated protein present in the disease state. The differences can be a result

of increases or decreases in ubiquitin ligation cascade activity or a result of increases or decreases in deubiquitination activity. Those of skill will recognize that ubiquitin maps are useful to determine whether a particular disease or disorder is associated with ubiquitin. Also ubiquitin maps can be used for diagnosis of a disorder associated with ubiquitin in an individual.

[0071] A "cancer cell" or a "transformed cell" is a cell that exhibits deregulation of proliferation. The deregulation can be a result of an increase in proliferation or a decrease in cell death, either apoptotic or necrotic cell death. Cancer cells and transformed cells can be derived from solid tumors or hematologic malignancies. Cancer cells and transformed cells also include transformed cell lines.

[0072] "Inhibitors", "activators", and "modulators" of ubiquitin ligation cascade polynucleotide and polypeptide sequences are used to refer to activating, inhibitory, or modulating molecules that inhibit, activate, or modulate *in vitro* and *in vivo* activity of ubiquitin ligation cascade polynucleotide and polypeptide sequences. Inhibitors are compounds that, *e.g.*, bind to, partially or totally block activity, decrease, prevent, delay activation, inactivate, desensitize, or down regulate the activity or expression of ubiquitin ligation cascade proteins, *e.g.*, antagonists. "Activators" are compounds that increase, open, activate, facilitate, enhance activation, sensitize, agonize, or up regulate ubiquitin ligation cascade protein activity, *e.g.*, agonists. Activators also include overexpressed ubiquitin ligation cascade proteins, *e.g.*, overexpression of a ubiquitin ligation cascade protein that is endogenously expressed in a cell or expression of a ubiquitin ligation cascade protein that is not normally expressed in a cell. Modulators encompasses both activators and inhibitors. Inhibitors, activators, or modulators also include genetically modified versions of ubiquitin ligation cascade proteins, *e.g.*, versions with altered activity, as well as naturally occurring and synthetic ligands, substrates, antagonists, agonists, antibodies, peptides, cyclic peptides, nucleic acids, antisense molecules, ribozymes, small chemical molecules and the like. Assays for ubiquitin ligation cascade proteins and inhibitors, activators and modulators thereof include, *e.g.*, expressing ubiquitin ligation cascade protein *in vitro*, in cells, or cell membranes, applying putative modulator compounds, and then determining the effects on activity. Assays for activity of ubiquitin liagation cascade proteins are found *e.g.*, in USSN 09/542497, filed April 4, 2000; USSN 09/826,312, filed April 3, 2001; USSN 10/108,767, filed March 26, 2002; and USSN 10/152,156, filed May 20, 2002.

[0073] "Inhibitors", "activators", and "modulators" of deubiquitination polynucleotide and polypeptide sequences are used to refer to activating, inhibitory, or modulating molecules that inhibit, activate, or modulate *in vitro* and *in vivo* activity of deubiquitination polynucleotide and polypeptide sequences. Inhibitors are compounds that, *e.g.*, bind to, partially or totally block activity, decrease, prevent, delay activation, inactivate, desensitize, or down regulate the activity or expression of deubiquitination proteins, *e.g.*, antagonists. "Activators" are compounds that increase, open, activate, facilitate, enhance activation, sensitize, agonize, or up regulate deubiquitination protein activity, *e.g.*, agonists. Activators also include overexpressed deubiquitination proteins, *e.g.*, overexpression of a deubiquitination protein that is endogenously expressed in a cell or expression of a deubiquitination protein that is not normally expressed in a cell. Modulators encompasses both activators and inhibitors. Inhibitors, activators, or modulators also include genetically modified versions of deubiquitination proteins, *e.g.*, versions with altered activity, as well as naturally occurring and synthetic ligands, substrates, antagonists, agonists, antibodies, peptides, cyclic peptides, nucleic acids, antisense molecules, ribozymes, small chemical molecules and the like. Assays for deubiquitination proteins and inhibitors, activators and modulators thereof include, *e.g.*, expressing deubiquitination protein *in vitro*, in cells, or cell membranes, applying putative modulator compounds, and then determining the effects on activity. Assays for activity of deubiquitination proteins are found *e.g.*, in USSN 10/232,759, filed August 30, 2002, and incorporated by reference herein in its entirety.

[0074] "Inhibitors", "activators", and "modulators" of ubiquitination pathways include inhibitors, activators, or modulators of ubiquitin ligation cascade protein and inhibitors, activators, or modulators deubiquitination proteins.

[0075] In a preferred embodiment, one or more components of the present invention comprise a tag. By "tag" is meant an attached molecule or molecules useful for the identification or isolation of the attached component. Components having a tag are referred to as "tag-X", wherein X is the component. For example, a ubiquitin comprising a tag is referred to herein as "tag-ubiquitin". Preferably, the tag is covalently bound to the attached component. When more than one component of a combination has a tag, the tags will be numbered for identification, for example "tag1-ubiquitin". Components may comprise more than one tag, in which case each tag will be numbered, for example "tag 1,2-ubiquitin". Preferred tags include, but are not limited to, a label, a partner of a binding pair, and a surface

substrate binding molecule. As will be evident to the skilled artisan, many molecules may find use as more than one type of tag, depending upon how the tag is used.

[0076] By "label" is meant a molecule that can be directly (i.e., a primary label) or indirectly (i.e., a secondary label) detected; for example a label can be visualized and/or measured or otherwise identified so that its presence or absence can be known. As will be appreciated by those in the art, the manner in which this is done will depend on the label. Preferred labels include, but are not limited to, fluorescent labels, label enzymes and radioisotopes.

[0077] By "fluorescent label" is meant any molecule that may be detected via its inherent fluorescent properties. Suitable fluorescent labels include, but are not limited to, fluorescein, rhodamine, tetramethylrhodamine, eosin, erythrosin, coumarin, methyl-coumarins, pyrene, Malacite green, stilbene, Lucifer Yellow, Cascade Blue™, Texas Red, IAEDANS, EDANS, BODIPY FL, LC Red 640, Cy 5, Cy 5.5, LC Red 705 and Oregon green. Suitable optical dyes are described in the 1996 Molecular Probes Handbook by Richard P. Haugland, hereby expressly incorporated by reference. Suitable fluorescent labels also include, but are not limited to, green fluorescent protein (GFP; Chalfie, et al., *Science* 263(5148):802-805 (Feb 11, 1994); and EGFP; Clontech - Genbank Accession Number U55762), blue fluorescent protein (BFP; 1. Quantum Biotechnologies, Inc. 1801 de Maisonneuve Blvd. West, 8th Floor, Montreal (Quebec) Canada H3H 1J9; 2. Stauber, R. H. Biotechniques 24(3):462-471 (1998); 3. Heim, R. and Tsien, R. Y. Curr. Biol. 6:178-182 (1996)), enhanced yellow fluorescent protein (EYFP; 1. Clontech Laboratories, Inc., 1020 East Meadow Circle, Palo Alto, CA 94303), luciferase (Ichiki, et al., *J. Immunol.* 150(12):5408-5417 (1993)), β-galactosidase (Nolan, et al., *Proc Natl Acad Sci USA* 85(8):2603-2607 (Apr 1988)) and Renilla WO 92/15673; WO 95/07463; WO 98/14605; WO 98/26277; WO 99/49019; U.S. patent 5,292,658; U.S. patent 5,418,155; U.S. patent 5,683,888; U.S. patent 5,741,668; U.S. patent 5,777,079; U.S. patent 5,804,387; U.S. patent 5,874,304; U.S. patent 5,876,995; and U.S. patent 5,925,558) All of the above-cited references are expressly incorporated herein by reference.

[0078] It is important to remember that ubiquitin is ligated to substrate protein by its terminal carboxyl group to a lysine residue, including lysine residues on other ubiquitin. Therefore, attachment of labels or other tags should not interfere with either of these active groups on the ubiquitin. Amino acids may be added to the sequence of protein, through

means well known in the art and described herein, for the express purpose of providing a point of attachment for a label. In a preferred embodiment, one or more amino acids are added to the sequence of a component for attaching a tag thereto, preferably a fluorescent label. In a preferred embodiment, the amino acid to which a fluorescent label is attached is Cysteine.

[0079] By "label enzyme" is meant an enzyme which may be reacted in the presence of a label enzyme substrate which produces a detectable product. Suitable label enzymes for use in the present invention include but are not limited to, horseradish peroxidase, alkaline phosphatase and glucose oxidase. Methods for the use of such substrates are well known in the art. The presence of the label enzyme is generally revealed through the enzyme's catalysis of a reaction with a label enzyme substrate, producing an identifiable product. Such products may be opaque, such as the reaction of horseradish peroxidase with tetramethyl benzedine, and may have a variety of colors. Other label enzyme substrates, such as Luminol (available from Pierce Chemical Co.), have been developed that produce fluorescent reaction products. Methods for identifying label enzymes with label enzyme substrates are well known in the art and many commercial kits are available. Examples and methods for the use of various label enzymes are described in Savage *et al.*, *Previews* 247:6-9 (1998), Young, *J. Virol. Methods* 24:227-236 (1989), which are each hereby incorporated by reference in their entirety.

[0080] By "radioisotope" is meant any radioactive molecule. Suitable radioisotopes for use in the invention include, but are not limited to ^{14}C, ^{3}H, ^{32}P, ^{33}P, ^{35}S, ^{125}I, and ^{131}I. The use of radioisotopes as labels is well known in the art.

[0081] In addition, labels may be indirectly detected, that is, the tag is a partner of a binding pair. By "partner of a binding pair" is meant one of a first and a second moiety, wherein said first and said second moiety have a specific binding affinity for each other. Suitable binding pairs for use in the invention include, but are not limited to, antigens/antibodies (for example, digoxigenin/anti-digoxigenin, dinitrophenyl (DNP)/anti-DNP, dansyl-X-anti-dansyl, Fluorescein/anti-fluorescein, lucifer yellow/anti-lucifer yellow, and rhodamine anti-rhodamine), biotin/avid (or biotin/streptavidin) and calmodulin binding protein (CBP)/calmodulin. Other suitable binding pairs include polypeptides such as the FLAG-peptide (Hopp *et al.*, *BioTechnology*, 6:1204-1210 (1988)); the KT3 epitope peptide (Martin *et al.*, *Science*, 255:192-194 (1992)); tubulin epitope peptide (Skinner *et al.*, *J. Biol.*

Chem., 266:15163-15166 (1991)); and the T7 gene 10 protein peptide tag (Lutz-Freyermuth *et al., Proc. Natl. Acad. Sci. USA*, 87:6393-6397 (1990)) and the antibodies each thereto. Generally, in a preferred embodiment, the smaller of the binding pair partners serves as the tag, as steric considerations in ubiquitin ligation may be important. As will be appreciated by those in the art, binding pair partners may be used in applications other than for labeling, as is further described below.

[0082] As will be appreciated by those in the art, a partner of one binding pair may also be a partner of another binding pair. For example, an antigen (first moiety) may bind to a first antibody (second moiety) which may, in turn, be an antigen for a second antibody (third moiety). It will be further appreciated that such a circumstance allows indirect binding of a first moiety and a third moiety via an intermediary second moiety that is a binding pair partner to each.

[0083] As will be appreciated by those in the art, a partner of a binding pair may comprise a label, as described above. It will further be appreciated that this allows for a tag to be indirectly labeled upon the binding of a binding partner comprising a label. Attaching a label to a tag which is a partner of a binding pair, as just described, is referred to herein as "indirect labeling".

[0084] By "surface substrate binding molecule" and grammatical equivalents thereof is meant a molecule have binding affinity for a specific surface substrate, which substrate is generally a member of a binding pair applied, incorporated or otherwise attached to a surface. Suitable surface substrate binding molecules and their surface substrates include, but are not limited to poly-histidine (poly-his) or poly-histidine-glycine (poly-his-gly) tags and Nickel substrate; the Glutathione-S Transferase tag and its antibody substrate (available from Pierce Chemical); the flu HA tag polypeptide and its antibody 12CA5 substrate (Field *et al., Mol. Cell. Biol.*, 8:2159-2165 (1988)); the c-myc tag and the 8F9, 3C7, 6E10, G4, B7 and 9E10 antibody substrates thereto (Evan *et al., Molecular and Cellular Biology*, 5:3610-3616 (1985)); and the Herpes Simplex virus glycoprotein D (gD) tag and its antibody substrate (Paborsky *et al., Protein Engineering*, 3(6):547-553 (1990)). In general, surface binding substrate molecules useful in the present invention include, but are not limited to, polyhistidine structures (His-tags) that bind nickel substrates, antigens that bind to surface substrates comprising antibody, haptens that bind to avidin substrate (*e.g.*, biotin) and CBP that binds to surface substrate comprising calmodulin.

[0085] Production of antibody-embedded substrates is well known; see Slinkin *et al.*, *Bioconj. Chem.* 2:342-348 (1991); Torchilin *et al.*, *supra*; Trubetskoy *et al.*, *Bioconj. Chem.* 3:323-327 (1992); King *et al.*, *Cancer Res.* 54:6176-6185 (1994); and Wilbur *et al.*, *Bioconjugate Chem.* 5:220-235 (1994) (all of which are hereby expressly incorporated by reference), and attachment of or production of proteins with antigens is described above.

[0086] Calmodulin-embedded substrates are commercially available, and production of proteins with CBP is described in Simcox *et al.*, *Strategies* 8:40-43 (1995), which is hereby incorporated by reference in its entirety.

[0087] As will be appreciated by those in the art, tag-components of the invention can be made in various ways, depending largely upon the form of the tag. Components of the invention and tags are preferably attached by a covalent bond.

[0088] The production of tag-polypeptides by recombinant means when the tag is also a polypeptide is described below. Production of FLAG-labeled proteins is well known in the art and kits for such production are commercially available (for example, from Kodak and Sigma). Methods for the production and use of FLAG-labeled proteins are found, for example, in Winston *et al.*, *Genes and Devel.* 13:270-283 (1999), incorporated herein in its entirety, as well as product handbooks provided with the above-mentioned kits.

[0089] Biotinylation of target molecules and substrates is well known, for example, a large number of biotinylation agents are known, including amine-reactive and thiol-reactive agents, for the biotinylation of proteins, nucleic acids, carbohydrates, carboxylic acids; see chapter 4, Molecular Probes Catalog, Haugland, 6th Ed. 1996, hereby incorporated by reference. A biotinylated substrate can be attached to a biotinylated component via avidin or streptavidin. Similarly, a large number of haptenylation reagents are also known (Id.).

[0090] Methods for labeling of proteins with radioisotopes are known in the art. For example, such methods are found in Ohta *et al.*, *Molec. Cell* 3:535-541 (1999), which is hereby incorporated by reference in its entirety.

[0091] Production of proteins having His-tags by recombinant means is well known, and kits for producing such proteins are commercially available. Such a kit and its use is described in the QIAexpress Handbook from Quiagen by Joanne Crowe *et al.*, hereby expressly incorporated by reference.

[0092] The functionalization of labels with chemically reactive groups such as thiols, amines, carboxyls, etc. is generally known in the art. In a preferred embodiment, the tag is functionalized to facilitate covalent attachment.

[0093] The covalent attachment of the tag may be either direct or via a linker. In one embodiment, the linker is a relatively short coupling moiety, that is used to attach the molecules. A coupling moiety may be synthesized directly onto a component of the invention, ubiquitin for example, and contains at least one functional group to facilitate attachment of the tag. Alternatively, the coupling moiety may have at least two functional groups, which are used to attach a functionalized component to a functionalized tag, for example. In an additional embodiment, the linker is a polymer. In this embodiment, covalent attachment is accomplished either directly, or through the use of coupling moieties from the component or tag to the polymer. In a preferred embodiment, the covalent attachment is direct, that is, no linker is used. In this embodiment, the component preferably contains a functional group such as a carboxylic acid which is used for direct attachment to the functionalized tag. It should be understood that the component and tag may be attached in a variety of ways, including those listed above. What is important is that manner of attachment does not significantly alter the functionality of the component. For example, in tag-ubiquitin, the tag should be attached in such a manner as to allow the ubiquitin to be covalently bound to other ubiquitin to form polyubiquitin chains. As will be appreciated by those in the art, the above description of covalent attachment of a label and ubiquitin applies equally to the attachment of virtually any two molecules of the present disclosure.

[0094] In a preferred embodiment, the tag is functionalized to facilitate covalent attachment, as is generally outlined above. Thus, a wide variety of tags are commercially available which contain functional groups, including, but not limited to, isothiocyanate groups, amino groups, haloacetyl groups, maleimides, succinimidyl esters, and sulfonyl halides, all of which may be used to covalently attach the tag to a second molecule, as is described herein. The choice of the functional group of the tag will depend on the site of attachment to either a linker, as outlined above or a component of the invention. Thus, for example, for direct linkage to a carboxylic acid group of a ubiquitin, amino modified or hydrazine modified tags will be used for coupling via carbodiimide chemistry, for example using 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDAC) as is known in the art (see Set 9 and Set 11 of the Molecular Probes Catalog, supra; see also the Pierce 1994 Catalog and Handbook, pages T-155 to T-200, both of which are hereby incorporated by reference). In

one embodiment, the carbodiimide is first attached to the tag, such as is commercially available for many of the tags described herein.

[0095] In a preferred embodiment, ubiquitin is in the form of tag-ubiquitin.

[0096] In a preferred embodiment, ubiquitin is in the form of tag-ubiquitin, wherein, tag is a partner of a binding pair. Preferably in this embodiment the tag is FLAG and the binding partner is anti-FLAG. In a second preferred embodiment, the tag is poly-histidine and the binding partner is a nickel substrate.

[0097] A "small organic molecule" refers to an organic molecule, either naturally occurring or synthetic, that has a molecular weight of more than about 50 daltons and less than about 2500 daltons, preferably less than about 2000 daltons, preferably between about 100 to about 1000 daltons, more preferably between about 200 to about 500 daltons.

[0098] "Biological sample" include sections of tissues such as biopsy and autopsy samples, and frozen sections taken for histologic purposes. Such samples include blood, sputum, tissue, cultured cells, e.g., primary cultures, explants, and transformed cells, stool, urine, etc. A biological sample is typically obtained from a eukaryotic organism, most preferably a mammal such as a primate e.g., chimpanzee or human; cow; dog; cat; a rodent, e.g., guinea pig, rat, mouse; rabbit; or a bird; reptile; or fish.

[0099] The terms "identical" or percent "identity," in the context of two or more nucleic acids or polypeptide sequences, refer to two or more sequences or subsequences that are the same or have a specified percentage of amino acid residues or nucleotides that are the same (i.e., about 60% identity, preferably 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or higher identity over a specified region, when compared and aligned for maximum correspondence over a comparison window or designated region) as measured using a BLAST or BLAST 2.0 sequence comparison algorithms with default parameters described below, or by manual alignment and visual inspection (*see, e.g.*, NCBI web site http://www.ncbi.nlm.nih.gov/BLAST/ or the like). Such sequences are then said to be "substantially identical." This definition also refers to, or may be applied to, the compliment of a test sequence. The definition also includes sequences that have deletions and/or additions, as well as those that have substitutions. As described below, the preferred algorithms can account for gaps and the like. Preferably, identity exists over a region that is at least about 25 amino acids or nucleotides in length, or more preferably over a region that is 50-100 amino acids or nucleotides in length.

[0100] For sequence comparison, typically one sequence acts as a reference sequence, to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are entered into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. Preferably, default program parameters can be used, or alternative parameters can be designated. The sequence comparison algorithm then calculates the percent sequence identities for the test sequences relative to the reference sequence, based on the program parameters.

[0101] A "comparison window", as used herein, includes reference to a segment of any one of the number of contiguous positions selected from the group consisting of from 20 to 600, usually about 50 to about 200, more usually about 100 to about 150 in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned. Methods of alignment of sequences for comparison are well-known in the art. Optimal alignment of sequences for comparison can be conducted, e.g., by the local homology algorithm of Smith & Waterman, *Adv. Appl. Math.* 2:482 (1981), by the homology alignment algorithm of Needleman & Wunsch, *J. Mol. Biol.* 48:443 (1970), by the search for similarity method of Pearson & Lipman, *Proc. Nat'l. Acad. Sci. USA* 85:2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), or by manual alignment and visual inspection (*see, e.g., Current Protocols in Molecular Biology* (Ausubel *et al.*, eds. 1995 supplement)).

[0102] A preferred example of algorithm that is suitable for determining percent sequence identity and sequence similarity are the BLAST and BLAST 2.0 algorithms, which are described in Altschul *et al., Nuc. Acids Res.* 25:3389-3402 (1977) and Altschul *et al., J. Mol. Biol.* 215:403-410 (1990), respectively. BLAST and BLAST 2.0 are used, with the parameters described herein, to determine percent sequence identity for the nucleic acids and proteins of the invention. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/). This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul *et al., supra*). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are extended in both directions along each sequence for as

far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always > 0) and N (penalty score for mismatching residues; always < 0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, M=5, N=-4 and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength of 3, and expectation (E) of 10, and the BLOSUM62 scoring matrix (*see* Henikoff & Henikoff, *Proc. Natl. Acad. Sci. USA* 89:10915 (1989)) alignments (B) of 50, expectation (E) of 10, M=5, N=-4, and a comparison of both strands.

[0103] "Nucleic acid" refers to deoxyribonucleotides or ribonucleotides and polymers thereof in either single- or double-stranded form, and complements thereof. The term encompasses nucleic acids containing known nucleotide analogs or modified backbone residues or linkages, which are synthetic, naturally occurring, and non-naturally occurring, which have similar binding properties as the reference nucleic acid, and which are metabolized in a manner similar to the reference nucleotides. Examples of such analogs include, without limitation, phosphorothioates, phosphoramidates, methyl phosphonates, chiral-methyl phosphonates, 2-O-methyl ribonucleotides, peptide-nucleic acids (PNAs).

[0104] Unless otherwise indicated, a particular nucleic acid sequence also implicitly encompasses conservatively modified variants thereof (e.g., degenerate codon substitutions) and complementary sequences, as well as the sequence explicitly indicated. Specifically, degenerate codon substitutions may be achieved by generating sequences in which the third position of one or more selected (or all) codons is substituted with mixed-base and/or deoxyinosine residues (Batzer *et al.*, *Nucleic Acid Res.* 19:5081 (1991); Ohtsuka *et al.*, *J. Biol. Chem.* 260:2605-2608 (1985); Rossolini *et al.*, *Mol. Cell. Probes* 8:91-98 (1994)). The term nucleic acid is used interchangeably with gene, cDNA, mRNA, oligonucleotide, and polynucleotide.

[0105] A particular nucleic acid sequence also implicitly encompasses "splice variants." Similarly, a particular protein encoded by a nucleic acid implicitly encompasses any protein encoded by a splice variant of that nucleic acid. "Splice variants," as the name suggests, are products of alternative splicing of a gene. After transcription, an initial nucleic acid transcript may be spliced such that different (alternate) nucleic acid splice products encode different polypeptides. Mechanisms for the production of splice variants vary, but include alternate splicing of exons. Alternate polypeptides derived from the same nucleic acid by read-through transcription are also encompassed by this definition. Any products of a splicing reaction, including recombinant forms of the splice products, are included in this definition. An example of potassium channel splice variants is discussed in Leicher, *et al.*, *J. Biol. Chem.* 273(52):35095-35101 (1998).

[0106] The terms "polypeptide," "peptide" and "protein" are used interchangeably herein to refer to a polymer of amino acid residues. The terms apply to amino acid polymers in which one or more amino acid residue is an artificial chemical mimetic of a corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers and non-naturally occurring amino acid polymer.

[0107] The term "amino acid" refers to naturally occurring and synthetic amino acids, as well as amino acid analogs and amino acid mimetics that function in a manner similar to the naturally occurring amino acids. Naturally occurring amino acids are those encoded by the genetic code, as well as those amino acids that are later modified, e.g., hydroxyproline, γ-carboxyglutamate, and O-phosphoserine. Amino acid analogs refers to compounds that have the same basic chemical structure as a naturally occurring amino acid, i.e., an α carbon that is bound to a hydrogen, a carboxyl group, an amino group, and an R group, e.g., homoserine, norleucine, methionine sulfoxide, methionine methyl sulfonium. Such analogs have modified R groups (e.g., norleucine) or modified peptide backbones, but retain the same basic chemical structure as a naturally occurring amino acid. Amino acid mimetics refers to chemical compounds that have a structure that is different from the general chemical structure of an amino acid, but that functions in a manner similar to a naturally occurring amino acid.

[0108] Amino acids may be referred to herein by either their commonly known three letter symbols or by the one-letter symbols recommended by the IUPAC-IUB Biochemical Nomenclature Commission. Nucleotides, likewise, may be referred to by their commonly accepted single-letter codes.

[0109] "Conservatively modified variants" applies to both amino acid and nucleic acid sequences. With respect to particular nucleic acid sequences, conservatively modified variants refers to those nucleic acids which encode identical or essentially identical amino acid sequences, or where the nucleic acid does not encode an amino acid sequence, to essentially identical sequences. Because of the degeneracy of the genetic code, a large number of functionally identical nucleic acids encode any given protein. For instance, the codons GCA, GCC, GCG and GCU all encode the amino acid alanine. Thus, at every position where an alanine is specified by a codon, the codon can be altered to any of the corresponding codons described without altering the encoded polypeptide. Such nucleic acid variations are "silent variations," which are one species of conservatively modified variations. Every nucleic acid sequence herein which encodes a polypeptide also describes every possible silent variation of the nucleic acid. One of skill will recognize that each codon in a nucleic acid (except AUG, which is ordinarily the only codon for methionine, and TGG, which is ordinarily the only codon for tryptophan) can be modified to yield a functionally identical molecule. Accordingly, each silent variation of a nucleic acid which encodes a polypeptide is implicit in each described sequence with respect to the expression product, but not with respect to actual probe sequences.

[0110] As to amino acid sequences, one of skill will recognize that individual substitutions, deletions or additions to a nucleic acid, peptide, polypeptide, or protein sequence which alters, adds or deletes a single amino acid or a small percentage of amino acids in the encoded sequence is a "conservatively modified variant" where the alteration results in the substitution of an amino acid with a chemically similar amino acid. Conservative substitution tables providing functionally similar amino acids are well known in the art. Such conservatively modified variants are in addition to and do not exclude polymorphic variants, interspecies homologs, and alleles of the invention.

[0111] The following eight groups each contain amino acids that are conservative substitutions for one another: 1) Alanine (A), Glycine (G); 2) Aspartic acid (D), Glutamic acid (E); 3) Asparagine (N), Glutamine (Q); 4) Arginine (R), Lysine (K); 5) Isoleucine (I), Leucine (L), Methionine (M), Valine (V); 6) Phenylalanine (F), Tyrosine (Y), Tryptophan (W); 7) Serine (S), Threonine (T); and 8) Cysteine (C), Methionine (M) (*see, e.g.*, Creighton, *Proteins* (1984)).

[0112] Macromolecular structures such as polypeptide structures can be described in terms of various levels of organization. For a general discussion of this organization, *see, e.g.*, Alberts *et al.*, *Molecular Biology of the Cell* (3rd ed., 1994) and Cantor and Schimmel, *Biophysical Chemistry Part I: The Conformation of Biological Macromolecules* (1980). "Primary structure" refers to the amino acid sequence of a particular peptide. "Secondary structure" refers to locally ordered, three dimensional structures within a polypeptide. These structures are commonly known as domains, e.g., enzymatic domains (RING, ligase), extracellular domains, transmembrane domains, pore domains, and cytoplasmic tail domains. Domains are portions of a polypeptide that form a compact unit of the polypeptide and are typically 15 to 350 amino acids long. Exemplary domains include domains with enzymatic activity, e.g., a ligase or RING domain. Typical domains are made up of sections of lesser organization such as stretches of β-sheet and α-helices. "Tertiary structure" refers to the complete three dimensional structure of a polypeptide monomer. "Quaternary structure" refers to the three dimensional structure formed by the noncovalent association of independent tertiary units. Anisotropic terms are also known as energy terms.

[0113] A "label" or a "detectable moiety" is a composition detectable by spectroscopic, photochemical, biochemical, immunochemical, chemical, or other physical means. For example, useful labels include ^{32}P, fluorescent dyes, electron-dense reagents, enzymes (e.g., as commonly used in an ELISA), biotin, digoxigenin, or haptens and proteins which can be made detectable, e.g., by incorporating a radiolabel into the peptide or used to detect antibodies specifically reactive with the peptide.

[0114] The term "recombinant" when used with reference, e.g., to a cell, or nucleic acid, protein, or vector, indicates that the cell, nucleic acid, protein or vector, has been modified by the introduction of a heterologous nucleic acid or protein or the alteration of a native nucleic acid or protein, or that the cell is derived from a cell so modified. Thus, for example, recombinant cells express genes that are not found within the native (non-recombinant) form of the cell or express native genes that are otherwise abnormally expressed, under expressed or not expressed at all.

[0115] The term "heterologous" when used with reference to portions of a nucleic acid indicates that the nucleic acid comprises two or more subsequences that are not found in the same relationship to each other in nature. For instance, the nucleic acid is typically recombinantly produced, having two or more sequences from unrelated genes arranged to

make a new functional nucleic acid, e.g., a promoter from one source and a coding region from another source. Similarly, a heterologous protein indicates that the protein comprises two or more subsequences that are not found in the same relationship to each other in nature (e.g., a fusion protein).

[0116] The phrase "stringent hybridization conditions" refers to conditions under which a probe will hybridize to its target subsequence, typically in a complex mixture of nucleic acids, but to no other sequences. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridize specifically at higher temperatures. An extensive guide to the hybridization of nucleic acids is found in Tijssen, *Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Probes*, "Overview of principles of hybridization and the strategy of nucleic acid assays" (1993). Generally, stringent conditions are selected to be about 5-10°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength pH. The T_m is the temperature (under defined ionic strength, pH, and nucleic concentration) at which 50% of the probes complementary to the target hybridize to the target sequence at equilibrium (as the target sequences are present in excess, at T_m, 50% of the probes are occupied at equilibrium). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. For selective or specific hybridization, a positive signal is at least two times background, preferably 10 times background hybridization. Exemplary stringent hybridization conditions can be as following: 50% formamide, 5x SSC, and 1% SDS, incubating at 42°C, or, 5x SSC, 1% SDS, incubating at 65°C, with wash in 0.2x SSC, and 0.1% SDS at 65°C.

[0117] Nucleic acids that do not hybridize to each other under stringent conditions are still substantially identical if the polypeptides which they encode are substantially identical. This occurs, for example, when a copy of a nucleic acid is created using the maximum codon degeneracy permitted by the genetic code. In such cases, the nucleic acids typically hybridize under moderately stringent hybridization conditions. Exemplary "moderately stringent hybridization conditions" include a hybridization in a buffer of 40% formamide, 1 M NaCl, 1% SDS at 37°C, and a wash in 1X SSC at 45°C. A positive hybridization is at least twice background. Those of ordinary skill will readily recognize that alternative hybridization and wash conditions can be utilized to provide conditions of similar stringency. Additional guidelines for determining hybridization parameters are provided in numerous reference, e.g., and *Current Protocols in Molecular Biology*, ed. Ausubel, *et al.*

[0118] For PCR, a temperature of about 36°C is typical for low stringency amplification, although annealing temperatures may vary between about 32°C and 48°C depending on primer length. For high stringency PCR amplification, a temperature of about 62°C is typical, although high stringency annealing temperatures can range from about 50°C to about 65°C, depending on the primer length and specificity. Typical cycle conditions for both high and low stringency amplifications include a denaturation phase of 90°C - 95°C for 30 sec - 2 min., an annealing phase lasting 30 sec. - 2 min., and an extension phase of about 72°C for 1 - 2 min. Protocols and guidelines for low and high stringency amplification reactions are provided, e.g., in Innis *et al.* (1990) *PCR Protocols, A Guide to Methods and Applications*, Academic Press, Inc. N.Y.).

[0119] "Antibody" refers to a polypeptide comprising a framework region from an immunoglobulin gene or fragments thereof that specifically binds and recognizes an antigen. The recognized immunoglobulin genes include the kappa, lambda, alpha, gamma, delta, epsilon, and mu constant region genes, as well as the myriad immunoglobulin variable region genes. Light chains are classified as either kappa or lambda. Heavy chains are classified as gamma, mu, alpha, delta, or epsilon, which in turn define the immunoglobulin classes, IgG, IgM, IgA, IgD and IgE, respectively. Typically, the antigen-binding region of an antibody will be most critical in specificity and affinity of binding.

[0120] An exemplary immunoglobulin (antibody) structural unit comprises a tetramer. Each tetramer is composed of two identical pairs of polypeptide chains, each pair having one "light" (about 25 kD) and one "heavy" chain (about 50-70 kD). The N-terminus of each chain defines a variable region of about 100 to 110 or more amino acids primarily responsible for antigen recognition. The terms variable light chain (V_L) and variable heavy chain (V_H) refer to these light and heavy chains respectively.

[0121] Antibodies exist, e.g., as intact immunoglobulins or as a number of well-characterized fragments produced by digestion with various peptidases. Thus, for example, pepsin digests an antibody below the disulfide linkages in the hinge region to produce $F(ab)'_2$, a dimer of Fab which itself is a light chain joined to V_H-C_H1 by a disulfide bond. The $F(ab)'_2$ may be reduced under mild conditions to break the disulfide linkage in the hinge region, thereby converting the $F(ab)'_2$ dimer into an Fab' monomer. The Fab' monomer is essentially Fab with part of the hinge region (*see Fundamental Immunology* (Paul ed., 3d ed. 1993). While various antibody fragments are defined in terms of the digestion of an intact

antibody, one of skill will appreciate that such fragments may be synthesized *de novo* either chemically or by using recombinant DNA methodology. Thus, the term antibody, as used herein, also includes antibody fragments either produced by the modification of whole antibodies, or those synthesized *de novo* using recombinant DNA methodologies (e.g., single chain Fv) or those identified using phage display libraries (*see, e.g.*, McCafferty *et al.*, *Nature* 348:552-554 (1990))

[0122] For preparation of antibodies, e.g., recombinant, monoclonal, or polyclonal antibodies, many technique known in the art can be used (*see, e.g.*, Kohler & Milstein, *Nature* 256:495-497 (1975); Kozbor *et al.*, *Immunology Today* 4: 72 (1983); Cole *et al.*, pp. 77-96 in *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc. (1985); Coligan, *Current Protocols in Immunology* (1991); Harlow & Lane, *Antibodies, A Laboratory Manual* (1988); and Goding, *Monoclonal Antibodies: Principles and Practice* (2d ed. 1986)). The genes encoding the heavy and light chains of an antibody of interest can be cloned from a cell, e.g., the genes encoding a monoclonal antibody can be cloned from a hybridoma and used to produce a recombinant monoclonal antibody. Gene libraries encoding heavy and light chains of monoclonal antibodies can also be made from hybridoma or plasma cells. Random combinations of the heavy and light chain gene products generate a large pool of antibodies with different antigenic specificity (*see, e.g.*, Kuby, *Immunology* (3rd ed. 1997)). Techniques for the production of single chain antibodies or recombinant antibodies (U.S. Patent 4,946,778, U.S. Patent No. 4,816,567) can be adapted to produce antibodies to polypeptides of this invention. Also, transgenic mice, or other organisms such as other mammals, may be used to express humanized or human antibodies (*see, e.g.*, U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, Marks *et al.*, *Bio/Technology* 10:779-783 (1992); Lonberg *et al.*, *Nature* 368:856-859 (1994); Morrison, *Nature* 368:812-13 (1994); Fishwild *et al.*, *Nature Biotechnology* 14:845-51 (1996); Neuberger, *Nature Biotechnology* 14:826 (1996); and Lonberg & Huszar, *Intern. Rev. Immunol.* 13:65-93 (1995)). Alternatively, phage display technology can be used to identify antibodies and heteromeric Fab fragments that specifically bind to selected antigens (*see, e.g.*, McCafferty *et al.*, *Nature* 348:552-554 (1990); Marks *et al.*, *Biotechnology* 10:779-783 (1992)). Antibodies can also be made bispecific, i.e., able to recognize two different antigens (*see, e.g.*, WO 93/08829, Traunecker *et al.*, *EMBO J.* 10:3655-3659 (1991); and Suresh *et al.*, *Methods in Enzymology* 121:210 (1986)). Antibodies can also be heteroconjugates, e.g., two

covalently joined antibodies, or immunotoxins (*see, e.g.*, U.S. Patent No. 4,676,980 , WO 91/00360; WO 92/200373; and EP 03089).

[0123] Methods for humanizing or primatizing non-human antibodies are well known in the art. Generally, a humanized antibody has one or more amino acid residues introduced into it from a source which is non-human. These non-human amino acid residues are often referred to as import residues, which are typically taken from an import variable domain. Humanization can be essentially performed following the method of Winter and co-workers (*see, e.g.*, Jones *et al., Nature* 321:522-525 (1986); Riechmann *et al., Nature* 332:323-327 (1988); Verhoeyen *et al., Science* 239:1534-1536 (1988) and Presta, *Curr. Op. Struct. Biol.* 2:593-596 (1992)), by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. Accordingly, such humanized antibodies are chimeric antibodies (U.S. Patent No. 4,816,567), wherein substantially less than an intact human variable domain has been substituted by the corresponding sequence from a non-human species. In practice, humanized antibodies are typically human antibodies in which some CDR residues and possibly some FR residues are substituted by residues from analogous sites in rodent antibodies.

[0124] A "chimeric antibody" is an antibody molecule in which (a) the constant region, or a portion thereof, is altered, replaced or exchanged so that the antigen binding site (variable region) is linked to a constant region of a different or altered class, effector function and/or species, or an entirely different molecule which confers new properties to the chimeric antibody, e.g., an enzyme, toxin, hormone, growth factor, drug, etc.; or (b) the variable region, or a portion thereof, is altered, replaced or exchanged with a variable region having a different or altered antigen specificity.

[0125] In one embodiment, the antibody is conjugated to an "effector" moiety. The effector moiety can be any number of molecules, including labeling moieties such as radioactive labels or fluorescent labels, or can be a therapeutic moiety. In one aspect the antibody modulates the activity of the protein.

[0126] The phrase "specifically (or selectively) binds" to an antibody or "specifically (or selectively) immunoreactive with," when referring to a protein or peptide, refers to a binding reaction that is determinative of the presence of the protein, often in a heterogeneous population of proteins and other biologics. Thus, under designated immunoassay conditions, the specified antibodies bind to a particular protein at least two times the background and

more typically more than 10 to 100 times background. Specific binding to an antibody under such conditions requires an antibody that is selected for its specificity for a particular protein. For example, polyclonal antibodies raised to a ubiquitin ligation cascade protein, polymorphic variants, alleles, orthologs, and conservatively modified variants, or splice variants, or portions thereof, can be selected to obtain only those polyclonal antibodies that are specifically immunoreactive with ubiquitin ligation cascade proteins and not with other proteins. This selection may be achieved by subtracting out antibodies that cross-react with other molecules. A variety of immunoassay formats may be used to select antibodies specifically immunoreactive with a particular protein. For example, solid-phase ELISA immunoassays are routinely used to select antibodies specifically immunoreactive with a protein (*see, e.g.*, Harlow & Lane, *Antibodies, A Laboratory Manual* (1988) for a description of immunoassay formats and conditions that can be used to determine specific immunoreactivity).

[0127] By "therapeutically effective dose" herein is meant a dose that produces effects for which it is administered. The exact dose will depend on the purpose of the treatment, and will be ascertainable by one skilled in the art using known techniques (*see, e.g.*, Lieberman, *Pharmaceutical Dosage Forms* (vols. 1-3, 1992); Lloyd, *The Art, Science and Technology of Pharmaceutical Compounding* (1999); and Pickar, *Dosage Calculations* (1999)).

IDENTIFICATION OF UBIQUITINATED PROTEINS

[0128] Identification of ubiquitinated proteins advances understanding of cellular processes that are regulated by proteolysis. Ubiquitin maps, as described herein, provide a list of ubiquitinated proteins under a given condition or in a particular cell type or subject. Ubiquitin maps can be used, for example, to identify previously unknown targets of ubiquitination; to determine the effects of modulators of the ubiquitin ligation pathway, to determine the effect of increased or decreased expression of a particular protein on the ubiquitin ligation pathway, *e.g.*, an E1, E2, or E3 protein; to compare ubiquitination in a diseased cell or subject versus a non-diseased cell or subject, *e.g.*, in cancer cells versus non-cancer cells; and to determine ubiquitination targets under particular cellular conditions, *e.g.*, during cell cycle arrest.

[0129] Ubiquitin maps identify the protein targets of ubiquitination in a cell or cell component. Ubiquitinated proteins are first detected and then separated from non-ubiquitinated proteins. Generally speaking, ubiquitinated proteins comprise polyubiquitin. A

variety of method can be used to detect ubiquitinated proteins and can be based on detection of either naturally occurring ubiquitin species or recombinant ubiquitin.

[0130] If naturally occurring ubiquitin is used, antibodies directed against polyubiquitin can be used to immunopurify the ubiquitinated proteins. In another embodiment, polyubiquitin can be detected by using proteins or protein domains that recognize ubiquitinated proteins. For example, the S5a subunit of the 26S proteosome binds polyubiquitinated proteins. (Young *et al., J. Biol. Chem.* 273: 5461-5467 (1998)). Ubiquitinated proteins can be purified using affinity chromatography with the S5a protein or a domain of the S5a protein.

[0131] Recombinant ubiquitin can also be used in the present invention. In one embodiments, cells of interest are transduced with a DNA vector the expresses ubiquitin.

[0132] In another embodiment, the recombinant ubiquitin is first purified from a convenient source, *e.g.*, bacteria. The cells of interest are lysed and then combined with the purified recombinant ubiquitin and incubated under conditions that will allow ubiquitination of target proteins to occur. For example, an ATP regenerating system can be included in the reaction.

[0133] In a preferred embodiment, the recombinant ubiquitin protein is labeled. In some embodiments, the label is a partner of a binding pair. By "partner of a binding pair" is meant one of a first and a second moiety, wherein said first and said second moiety have a specific binding affinity for each other. Suitable binding pairs for use in the invention include, but are not limited to, antigens/antibodies (for example, digoxigenin/anti-digoxigenin, dinitrophenyl (DNP)/anti-DNP, dansyl-X-anti-dansyl, Fluorescein/anti-fluorescein, lucifer yellow/anti-lucifer yellow, and rhodamine anti-rhodamine), biotin/avid (or biotin/streptavidin) and calmodulin binding protein (CBP)/calmodulin. Other suitable binding pairs include polypeptides such as the FLAG-peptide, the KT3 epitope peptide, tubulin epitope peptide, and the T7 gene 10 protein peptide tag, and the antibodies each thereto. (*See e.g.*, Hopp *et al.*, *BioTechnology*, 6:1204-1210 (1988); Martin *et al.*, *Science*, 255:192-194 (1992); Skinner *et al.*, *J. Biol. Chem.*, 266:15163-15166 (1991); and Lutz-Freyermuth *et al.*, *Proc. Natl. Acad. Sci. USA*, 87:6393-6397 (1990)). Generally, in a preferred embodiment, the smaller of the binding pair partners serves as the label, as steric considerations in ubiquitin ligation may be important.

[0134] In other embodiments the label is a surface substrate binding molecule. Suitable surface substrate binding molecules and their surface substrates include, but are not limited to

poly-histidine (poly-his) or poly-histidine-glycine (poly-his-gly) tags and Nickel substrate; the Glutathione-S Transferase tag and its antibody substrate (available from Pierce Chemical), the flu HA tag polypeptide and its antibody 12CA5 substrate, the c-myc tag and the 8F9, 3C7, 6E10, G4, B7 and 9E10 antibody substrates thereto, and the Herpes Simplex virus glycoprotein D (gD) tag and its antibody substrate. (*See, e.g.*, Field *et al.*, *Mol. Cell. Biol.*, 8:2159-2165 (1988); Evan *et al.*, *Molecular and Cellular Biology*, 5:3610-3616 (1985); Paborsky *et al.*, *Protein Engineering*, 3(6):547-553 (1990)). In general, surface binding substrate molecules useful in the present invention include, but are not limited to, polyhistidine structures (His-tags) that bind nickel substrates, antigens that bind to surface substrates comprising antibody, haptens that bind to avidin substrate (*e.g.*, biotin) and CBP that binds to surface substrate comprising calmodulin.

[0135] Those of skill will recognize steps that can be taken to increase the yield of ubiquitinated substrate proteins, if necessary. For example, for some target proteins it can be beneficial to first isolate a subcellular fraction and then isolate ubiquitinated proteins from that fraction. Also, cells can be incubated with inhibitors of downstream processing of ubiquitinated proteins, *e.g.*, proteasome inhibitors, such as MG132, PS341 and epoxymyosin; and deubiquitinase inhibitors, such as ubiquitin aldehyde.

[0136] In some embodiments, the ubiquitinated proteins are cleaved before further analysis occurs. The proteins can be cleaved by a specific protease, for example, lys C endoprotease or trypsin.

[0137] Once ubiquitinated proteins have been separated from non-ubiquitinated proteins, the ubiquitinated proteins or peptides derived from the ubiquitinated proteins are isolated from each other. Those of skill are aware that many method can be used to isolate proteins. For example, column chromatography can be used to isolate proteins from each other. For example, separation can be based on differences in size, net surface charge, hydrophobicity and affinity for ligands or substrates of the ubiquitinated proteins. In some embodiments, a ubiquitinated protein will be purified using conventional purification techniques before further analysis takes place, *i.e.*, LC-MS/MS.

[0138] Liquid chromatography systems can be used to isolate ubiquitinated proteins from each other. In a preferred embodiment, the ubiquitinated proteins are first separated using and ion exchange resin in an HPLC system, followed by separation using reverse phase media and buffer system. Those of skill will recognize that it can be advantageous to "desalt"

a sample before applying to an ion exchange column. The isolated proteins or peptides from the HPLC column can be processed immediately by a mass spectrometer.

[0139] Two dimensional (2D) gel electrophoresis can also be used to isolate ubiquitinated proteins from each other. In 2D gel electrophoresis proteins are separated in a polyacrylamide gel on the basis of charge (the proteins' isoelectric point) in the first dimension and by mass in the second dimension. 2D gels can be run using an equilibrium pH gradient or a non-equilibrium pH gradient (NEPHGE). Gels are stained to visualize the isolated proteins. Proteins of interest are excised and subjected to further analysis, *e.g.*, cleavage by proteases, liquid chromatography, or mass spectroscopy.

[0140] Briefly, in mass spectroscopy (MS) a protein or peptide sample is ionized and accelerated through a magnetic field to the a detector. Peptides of different mass and charge follow different paths through the instrument, resulting in different times of flight and producing a series of peaks corresponding to each peptide fragment. The peptide fingerprint is then matched against a protein database. (See *e.g.*, Mitchell, *Nature Biotech.* 21:233-237 (2003)).

[0141] Samples can be introduced into a mass spectrometer using a variety of methods, *e.g.*, direct insertion probes, direct infusion, or a capillary column. A number of ionization techniques can be used to ionize the sample, *e.g.*, fast atom/ion bombardment (FAB), matrix-assisted laser desorption/ionization (MALDI), and electrospray ionization (ESI). An number of mass analyzers can be used, *e.g.*, quadrupole analyzers, ion traps, magnetic mass analyzers, double focusing magnetic sector mass analyzer, time-of-flight (TOF) analyzer, Fourier transform-ion cyclotron resonance. Tandem mass spectrometry (MS/MS) can also be used, *e.g.*, MS/MS with a triple quadrupole mass spectrometer, MS/MS with a two-sector mass spectrometer, MSMS with a four-sector mass spectrometer, MS/MS with a two-sector quadrupole mass spectrometer, MS/MS with a time-of-flight reflectron mass spectrometer, MS/MS with a Fourier transform-ion cyclotron resonance, MS/MS with an ion trap, and TOF/TOF. A number of ion detectors can be used including electron multipliers and scintillation counters.

[0142] In a preferred embodiment, a Q-TOF tandem mass spectrometer is used for analysis of the isolated ubiquitinated proteins.

[0143] Once collected, the MS or MS/MS data is compared to databases of known proteins and used to identify the ubiquitinated proteins. Before conducting the database search using

the collected MS/MS data, the raw data can be pre-filtered to eliminate spectra found in control samples. For example, a CGI Perl program called "MS2Filter" can be used to pre-filter the raw data. This program digitally compares all experimental MS/MS spectra in quadruplicate runs with those found in each of the quadruplicate control runs and eliminates them before initiating the database search. This pre-filtering of the MS/MS data significantly reduced the bulk of the MS/MS data and increased the speed of protein identification. After pre-filtering, the remaining mass spectra are analyzed against a human protein sequence database. In a preferred embodiment, TurboSequest software (Thermofinnigan, San Jose, CA), which uses the SEQUEST algorithm, against a human protein sequence database derived from the NCBI non redundant human database (INFORMATICS Websites used: ncbi.nlm.nih.gov:80/entrez; isrec.isbsib.ch/software/PFSCAN_form.html; blocks.fhcrc.org/blocks/blocks_search.html) is used. If appropriate, the data can be further analyzed. In a preferred embodiment, SEQUEST results are summarized and further analyzed using a separate database called "MEDUSA" (Gururaja *et al., J. Proteome Res.* 1:253-61 (2002)). MEDUSA is a proprietary web-based Oracle 8.0 database that we have designed and coded that stores and evaluates mass spectrometry data and SEQUEST output. It runs independently of SEQUEST.

ISOLATION OF NUCLEIC ACIDS

[0144] This invention relies on routine techniques in the field of recombinant genetics. Basic texts disclosing the general methods of use in this invention include Sambrook *et al.*, *Molecular Cloning, A Laboratory Manual* (2nd ed. 1989); Kriegler, *Gene Transfer and Expression: A Laboratory Manual* (1990); and *Current Protocols in Molecular Biology* (Ausubel *et al.*, eds., 1994)).

[0145] Ubiquitin ligation cascade nucleic acids, polymorphic variants, orthologs, and alleles can be isolated using nucleic acid probes and oligonucleotides under stringent hybridization conditions, by screening libraries. Alternatively, expression libraries can be used to clone the ubiquitin ligation cascade protein, polymorphic variants, orthologs, and alleles by detecting expressed homologs immunologically with antisera or purified antibodies made against human ubiquitin ligation cascade proteins or portions thereof.

[0146] To make a cDNA library, one should choose a source that is rich in the RNA of choice. The mRNA is then made into cDNA using reverse transcriptase, ligated into a recombinant vector, and transfected into a recombinant host for propagation, screening and

cloning. Methods for making and screening cDNA libraries are well known (*see, e.g.*, Gubler & Hoffman, *Gene* 25:263-269 (1983); Sambrook *et al., supra*; Ausubel *et al., supra*).

[0147] For a genomic library, the DNA is extracted from the tissue and either mechanically sheared or enzymatically digested to yield fragments of about 12-20 kb. The fragments are then separated by gradient centrifugation from undesired sizes and are constructed in bacteriophage lambda vectors. These vectors and phage are packaged *in vitro*. Recombinant phage are analyzed by plaque hybridization as described in Benton & Davis, *Science* 196:180-182 (1977). Colony hybridization is carried out as generally described in Grunstein *et al., Proc. Natl. Acad. Sci. USA.*, 72:3961-3965 (1975).

[0148] A preferred method of isolating ubiquitin ligation cascade nucleic acids and orthologs, alleles, mutants, polymorphic variants, splice variants, and conservatively modified variants combines the use of synthetic oligonucleotide primers and amplification of an RNA or DNA template (*see* U.S. Patents 4,683,195 and 4,683,202; *PCR Protocols: A Guide to Methods and Applications* (Innis *et al.*, eds, 1990)). Methods such as polymerase chain reaction (PCR and RT-PCR) and ligase chain reaction (LCR) can be used to amplify nucleic acid sequences directly from mRNA, from cDNA, from genomic libraries or cDNA libraries. Degenerate oligonucleotides can be designed to amplify homologs using the sequences provided herein. Restriction endonuclease sites can be incorporated into the primers. Polymerase chain reaction or other *in vitro* amplification methods may also be useful, for example, to clone nucleic acid sequences that code for proteins to be expressed, to make nucleic acids to use as probes for detecting the presence of ubiquitin ligation cascade component encoding mRNA in physiological samples, for nucleic acid sequencing, or for other purposes. Genes amplified by the PCR reaction can be purified from agarose gels and cloned into an appropriate vector.

[0149] Gene expression of ubiquitin ligation cascade components can also be analyzed by techniques known in the art, e.g., reverse transcription and amplification of mRNA, isolation of total RNA or poly A^+ RNA, northern blotting, dot blotting, *in situ* hybridization, RNase protection, high density polynucleotide array technology, e.g., and the like.

[0150] Nucleic acids encoding ubiquitin ligation cascade proteins can be used with high density oligonucleotide array technology (e.g., GeneChip™) to identify ubiquitin ligation cascade proteins, orthologs, alleles, conservatively modified variants, and polymorphic variants in this invention. In the case where the homologs being identified are linked to

modulation of the cell cycle, they can be used with GeneChip™ as a diagnostic tool in detecting the disease in a biological sample, *see, e.g.*, Gunthand *et al., AIDS Res. Hum. Retroviruses* 14: 869-876 (1998); Kozal *et al., Nat. Med.* 2:753-759 (1996); Matson *et al., Anal. Biochem.* 224:110-106 (1995); Lockhart *et al., Nat. Biotechnol.* 14:1675-1680 (1996); Gingeras *et al., Genome Res.* 8:435-448 (1998); Hacia *et al., Nucleic Acids Res.* 26:3865-3866 (1998).

[0151] The gene of choice is typically cloned into intermediate vectors before transformation into prokaryotic or eukaryotic cells for replication and/or expression. These intermediate vectors are typically prokaryote vectors, e.g., plasmids, or shuttle vectors.

EXPRESSION IN PROKARYOTES AND EUKARYOTES

[0152] To obtain high level expression of a cloned gene, such as those cDNAs encoding ubiquitin ligation cascade proteins, one typically subclones the nucleic acid into an expression vector that contains a strong promoter to direct transcription, a transcription/translation terminator, and if for a nucleic acid encoding a protein, a ribosome binding site for translational initiation. Suitable bacterial promoters are well known in the art and described, e.g., in Sambrook *et al.*, and Ausubel *et al, supra*. Bacterial expression systems for expressing the protein are available in, e.g., *E. coli, Bacillus sp.,* and *Salmonella* (Palva *et al.*, *Gene* 22:229-235 (1983); Mosbach *et al.*, *Nature* 302:543-545 (1983). Kits for such expression systems are commercially available. Eukaryotic expression systems for mammalian cells, yeast, and insect cells are well known in the art and are also commercially available. In one preferred embodiment, retroviral expression systems are used in the present invention.

[0153] Selection of the promoter used to direct expression of a heterologous nucleic acid depends on the particular application. The promoter is preferably positioned about the same distance from the heterologous transcription start site as it is from the transcription start site in its natural setting. As is known in the art, however, some variation in this distance can be accommodated without loss of promoter function.

[0154] In addition to the promoter, the expression vector typically contains a transcription unit or expression cassette that contains all the additional elements required for the expression of the nucleic acid in host cells. A typical expression cassette thus contains a promoter operably linked to the nucleic acid sequence encoding the nucleic acid of choice and signals required for efficient polyadenylation of the transcript, ribosome binding sites,

and translation termination. Additional elements of the cassette may include enhancers and, if genomic DNA is used as the structural gene, introns with functional splice donor and acceptor sites.

[0155] In addition to a promoter sequence, the expression cassette should also contain a transcription termination region downstream of the structural gene to provide for efficient termination. The termination region may be obtained from the same gene as the promoter sequence or may be obtained from different genes.

[0156] The particular expression vector used to transport the genetic information into the cell is not particularly critical. Any of the conventional vectors used for expression in eukaryotic or prokaryotic cells may be used. Standard bacterial expression vectors include plasmids such as pBR322 based plasmids, pSKF, pET23D, and fusion expression systems such as MBP, GST, and LacZ. Epitope tags can also be added to recombinant proteins to provide convenient methods of isolation, e.g., c-myc. Sequence tags may be included in an expression cassette for nucleic acid rescue. Markers such as fluorescent proteins, green or red fluorescent protein, β-gal, CAT, and the like can be included in the vectors as markers for vector transduction.

[0157] Expression vectors containing regulatory elements from eukaryotic viruses are typically used in eukaryotic expression vectors, e.g., SV40 vectors, papilloma virus vectors, retroviral vectors, and vectors derived from Epstein-Barr virus. Other exemplary eukaryotic vectors include pMSG, pAV009/A$^+$, pMTO10/A$^+$, pMAMneo-5, baculovirus pDSVE, and any other vector allowing expression of proteins under the direction of the CMV promoter, SV40 early promoter, SV40 later promoter, metallothionein promoter, murine mammary tumor virus promoter, Rous sarcoma virus promoter, polyhedrin promoter, or other promoters shown effective for expression in eukaryotic cells.

[0158] Expression of proteins from eukaryotic vectors can be also be regulated using inducible promoters. With inducible promoters, expression levels are tied to the concentration of inducing agents, such as tetracycline or ecdysone, by the incorporation of response elements for these agents into the promoter. Generally, high level expression is obtained from inducible promoters only in the presence of the inducing agent; basal expression levels are minimal.

[0159] In one embodiment, the vectors of the invention have a regulatable promoter, e.g., tet-regulated systems and the RU-486 system (*see, e.g.*, Gossen & Bujard, *PNAS* 89:5547

(1992); Oligino *et al., Gene Ther.* 5:491-496 (1998); Wang *et al., Gene Ther.* 4:432-441 (1997); Neering *et al., Blood* 88:1147-1155 (1996); and Rendahl *et al., Nat. Biotechnol.* 16:757-761 (1998)). These impart small molecule control on the expression of the candidate target nucleic acids. This beneficial feature can be used to determine that a desired phenotype is caused by a transfected cDNA rather than a somatic mutation.

[0160] Some expression systems have markers that provide gene amplification such as thymidine kinase and dihydrofolate reductase. Alternatively, high yield expression systems not involving gene amplification are also suitable, such as using a baculovirus vector in insect cells, with a sequence of choice under the direction of the polyhedrin promoter or other strong baculovirus promoters.

[0161] The elements that are typically included in expression vectors also include a replicon that functions in *E. coli*, a gene encoding antibiotic resistance to permit selection of bacteria that harbor recombinant plasmids, and unique restriction sites in nonessential regions of the plasmid to allow insertion of eukaryotic sequences. The particular antibiotic resistance gene chosen is not critical, any of the many resistance genes known in the art are suitable. The prokaryotic sequences are preferably chosen such that they do not interfere with the replication of the DNA in eukaryotic cells, if necessary.

[0162] Standard transfection methods are used to produce bacterial, mammalian, yeast or insect cell lines that express large quantities of protein, which are then purified using standard techniques (*see, e.g.*, Colley *et al., J. Biol. Chem.* 264:17619-17622 (1989); *Guide to Protein Purification*, in *Methods in Enzymology*, vol. 182 (Deutscher, ed., 1990)). Transformation of eukaryotic and prokaryotic cells are performed according to standard techniques (*see, e.g.*, Morrison, *J. Bact.* 132:349-351 (1977); Clark-Curtiss & Curtiss, *Methods in Enzymology* 101:347-362 (Wu *et al.*, eds, 1983).

[0163] Any of the well-known procedures for introducing foreign nucleotide sequences into host cells may be used. These include the use of calcium phosphate transfection, polybrene, protoplast fusion, electroporation, biolistics, liposomes, microinjection, plasma vectors, viral vectors and any of the other well known methods for introducing cloned genomic DNA, cDNA, synthetic DNA or other foreign genetic material into a host cell (*see, e.g.*, Sambrook *et al., supra*). It is only necessary that the particular genetic engineering procedure used be capable of successfully introducing at least one gene into the host cell capable of expressing ubiquitin ligation cascade components.

[0164] After the expression vector is introduced into the cells, the transfected cells are cultured under conditions favoring expression of the protein of choice, which is recovered from the culture using standard techniques identified below.

[0165] **PURIFICATION OF POLYPEPTIDES**

[0166] Either naturally occurring or recombinant ubiquitin ligation cascade components can be purified for use in functional assays. Naturally occurring protein can be purified, e.g., from human tissue. Recombinant protein can be purified from any suitable expression system.

[0167] The protein may be purified to substantial purity by standard techniques, including selective precipitation with such substances as ammonium sulfate; column chromatography, immunopurification methods, and others (*see, e.g.*, Scopes, *Protein Purification: Principles and Practice* (1982); U.S. Patent No. 4,673,641; Ausubel *et al., supra;* and Sambrook *et al., supra*).

[0168] A number of procedures can be employed when recombinant protein is being purified. For example, proteins having established molecular adhesion properties can be reversible fused to the protein. With the appropriate ligand or substrate, a specific protein can be selectively adsorbed to a purification column and then freed from the column in a relatively pure form. The fused protein is then removed by enzymatic activity. Finally, protein could be purified using immunoaffinity columns. Recombinant protein can be purified from any suitable source, include yeast, insect, bacterial, and mammalian cells.

A. Purification of protein from recombinant bacteria

[0169] Recombinant proteins are expressed by transformed bacteria in large amounts, typically after promoter induction; but expression can be constitutive. Promoter induction with IPTG is one example of an inducible promoter system. Bacteria are grown according to standard procedures in the art. Fresh or frozen bacteria cells are used for isolation of protein.

[0170] Proteins expressed in bacteria may form insoluble aggregates ("inclusion bodies"). Several protocols are suitable for purification of protein inclusion bodies. For example, purification of inclusion bodies typically involves the extraction, separation and/or purification of inclusion bodies by disruption of bacterial cells, e.g., by incubation in a buffer of 50 mM TRIS/HCL pH 7.5, 50 mM NaCl, 5 mM $MgCl_2$, 1 mM DTT, 0.1 mM ATP, and 1 mM PMSF. The cell suspension can be lysed using 2-3 passages through a French Press,

homogenized using a Polytron (Brinkman Instruments) or sonicated on ice. Alternate methods of lysing bacteria are apparent to those of skill in the art (*see, e.g.*, Sambrook *et al.*, *supra*; Ausubel *et al.*, *supra*).

[0171] If necessary, the inclusion bodies are solubilized, and the lysed cell suspension is typically centrifuged to remove unwanted insoluble matter. Proteins that formed the inclusion bodies may be renatured by dilution or dialysis with a compatible buffer. Suitable solvents include, but are not limited to urea (from about 4 M to about 8 M), formamide (at least about 80%, volume/volume basis), and guanidine hydrochloride (from about 4 M to about 8 M). Some solvents which are capable of solubilizing aggregate-forming proteins, for example SDS (sodium dodecyl sulfate), 70% formic acid, are inappropriate for use in this procedure due to the possibility of irreversible denaturation of the proteins, accompanied by a lack of immunogenicity and/or activity. Although guanidine hydrochloride and similar agents are denaturants, this denaturation is not irreversible and renaturation may occur upon removal (by dialysis, for example) or dilution of the denaturant, allowing re-formation of immunologically and/or biologically active protein. Other suitable buffers are known to those skilled in the art. Human proteins are separated from other bacterial proteins by standard separation techniques, e.g., with Ni-NTA agarose resin.

[0172] Alternatively, it is possible to purify recombinant protein from bacteria periplasm. After lysis of the bacteria, the periplasmic fraction of the bacteria can be isolated by cold osmotic shock in addition to other methods known to skill in the art. To isolate recombinant proteins from the periplasm, the bacterial cells are centrifuged to form a pellet. The pellet is resuspended in a buffer containing 20% sucrose. To lyse the cells, the bacteria are centrifuged and the pellet is resuspended in ice-cold 5 mM $MgSO_4$ and kept in an ice bath for approximately 10 minutes. The cell suspension is centrifuged and the supernatant decanted and saved. The recombinant proteins present in the supernatant can be separated from the host proteins by standard separation techniques well known to those of skill in the art.

B. Standard protein separation techniques for purifying proteins

Solubility fractionation

[0173] Often as an initial step, particularly if the protein mixture is complex, an initial salt fractionation can separate many of the unwanted host cell proteins (or proteins derived from the cell culture media) from the recombinant protein of interest. The preferred salt is ammonium sulfate. Ammonium sulfate precipitates proteins by effectively reducing the

amount of water in the protein mixture. Proteins then precipitate on the basis of their solubility. The more hydrophobic a protein is, the more likely it is to precipitate at lower ammonium sulfate concentrations. A typical protocol includes adding saturated ammonium sulfate to a protein solution so that the resultant ammonium sulfate concentration is between 20-30%. This concentration will precipitate the most hydrophobic of proteins. The precipitate is then discarded (unless the protein of interest is hydrophobic) and ammonium sulfate is added to the supernatant to a concentration known to precipitate the protein of interest. The precipitate is then solubilized in buffer and the excess salt removed if necessary, either through dialysis or diafiltration. Other methods that rely on solubility of proteins, such as cold ethanol precipitation, are well known to those of skill in the art and can be used to fractionate complex protein mixtures.

Size differential filtration

[0174] The molecular weight of the protein can be used to isolate it from proteins of greater and lesser size using ultrafiltration through membranes of different pore size (for example, Amicon or Millipore membranes). As a first step, the protein mixture is ultrafiltered through a membrane with a pore size that has a lower molecular weight cut-off than the molecular weight of the protein of interest. The retentate of the ultrafiltration is then ultrafiltered against a membrane with a molecular cut off greater than the molecular weight of the protein of interest. The recombinant protein will pass through the membrane into the filtrate. The filtrate can then be chromatographed as described below.

Column chromatography

[0175] The protein can also be separated from other proteins on the basis of its size, net surface charge, hydrophobicity, and affinity for ligands or substrates. In addition, antibodies raised against proteins can be conjugated to column matrices and the proteins immunopurified. All of these methods are well known in the art. It will be apparent to one of skill that chromatographic techniques can be performed at any scale and using equipment from many different manufacturers (e.g., Pharmacia Biotech).

IMMUNOLOGICAL DETECTION OF POLYPEPTIDES

[0176] In addition to the detection of a ubiquitin ligation cascade gene and gene expression using nucleic acid hybridization technology, one can also use immunoassays to detect ubiquitin ligation cascade proteins of the invention. Such assays are useful for screening for modulators of ubiquitin ligation cascade proteins, as well as for therapeutic and diagnostic

applications. Immunoassays can be used to qualitatively or quantitatively analyze ubiquitin ligation cascade protein. A general overview of the applicable technology can be found in Harlow & Lane, *Antibodies: A Laboratory Manual* (1988).

A. Production of antibodies

[0177] Methods of producing polyclonal and monoclonal antibodies that react specifically with ubiquitin ligation cascade proteins are known to those of skill in the art (*see, e.g.*, Coligan, *Current Protocols in Immunology* (1991); Harlow & Lane, *supra*; Goding, *Monoclonal Antibodies: Principles and Practice* (2d ed. 1986); and Kohler & Milstein, *Nature* 256:495-497 (1975). Such techniques include antibody preparation by selection of antibodies from libraries of recombinant antibodies in phage or similar vectors, as well as preparation of polyclonal and monoclonal antibodies by immunizing rabbits or mice (*see, e.g.*, Huse *et al., Science* 246:1275-1281 (1989); Ward *et al., Nature* 341:544-546 (1989)).

[0178] A number of immunogens comprising portions of a ubiquitin ligation cascade protein may be used to produce antibodies specifically reactive with the ubiquitin ligation cascade protein. For example, a recombinant ubiquitin ligation cascade protein or an antigenic fragment thereof, can be isolated as described herein. Recombinant protein can be expressed in eukaryotic or prokaryotic cells as described above, and purified as generally described above. Recombinant protein is the preferred immunogen for the production of monoclonal or polyclonal antibodies. Alternatively, a synthetic peptide derived from the sequences disclosed herein and conjugated to a carrier protein can be used an immunogen. Naturally occurring protein may also be used either in pure or impure form. The product is then injected into an animal capable of producing antibodies. Either monoclonal or polyclonal antibodies may be generated, for subsequent use in immunoassays to measure the protein.

[0179] Methods of production of polyclonal antibodies are known to those of skill in the art. An inbred strain of mice (e.g., BALB/C mice) or rabbits is immunized with the protein using a standard adjuvant, such as Freund's adjuvant, and a standard immunization protocol. The animal's immune response to the immunogen preparation is monitored by taking test bleeds and determining the titer of reactivity to the beta subunits. When appropriately high titers of antibody to the immunogen are obtained, blood is collected from the animal and antisera are prepared. Further fractionation of the antisera to enrich for antibodies reactive to the protein can be done if desired (*see,* Harlow & Lane, *supra*).

[0180] Monoclonal antibodies may be obtained by various techniques familiar to those skilled in the art. Briefly, spleen cells from an animal immunized with a desired antigen are immortalized, commonly by fusion with a myeloma cell (*see,* Kohler & Milstein, *Eur. J. Immunol.* 6:511-519 (1976)). Alternative methods of immortalization include transformation with Epstein Barr Virus, oncogenes, or retroviruses, or other methods well known in the art. Colonies arising from single immortalized cells are screened for production of antibodies of the desired specificity and affinity for the antigen, and yield of the monoclonal antibodies produced by such cells may be enhanced by various techniques, including injection into the peritoneal cavity of a vertebrate host. Alternatively, one may isolate DNA sequences which encode a monoclonal antibody or a binding fragment thereof by screening a DNA library from human B cells according to the general protocol outlined by Huse, *et al.*, *Science* 246:1275-1281 (1989).

[0181] Monoclonal antibodies and polyclonal sera are collected and titered against the immunogen protein in an immunoassay, for example, a solid phase immunoassay with the immunogen immobilized on a solid support. Typically, polyclonal antisera with a titer of 10^4 or greater are selected and tested for their cross reactivity against non- ubiquitin ligation cascade proteins, using a competitive binding immunoassay. Specific polyclonal antisera and monoclonal antibodies will usually bind with a K_d of at least about 0.1 mM, more usually at least about 1 μM, preferably at least about 0.1 μM or better, and most preferably, 0.01 μM or better. Antibodies specific only for a particular ubiquitin ligation cascade protein ortholog, such as a human ortholog, can also be made, by subtracting out other cross-reacting orthologs from a species such as a non-human mammal. In this manner, antibodies that bind only to the protein of choice may be obtained.

[0182] Once the specific antibodies against a ubiquitin ligation cascade protein are available, the protein can be detected by a variety of immunoassay methods. In addition, the antibody can be used therapeutically as a ubiquitin ligation cascade protein modulators. For a review of immunological and immunoassay procedures, see *Basic and Clinical Immunology* (Stites & Terr eds., 7th ed. 1991). Moreover, the immunoassays of the present invention can be performed in any of several configurations, which are reviewed extensively in Enzyme Immunoassay (Maggio, ed., 1980); and Harlow & Lane, *supra.*

B. Immunological binding assays

[0183] Protein can be detected and/or quantified using any of a number of well recognized immunological binding assays (*see, e.g.,* U.S. Patents 4,366,241; 4,376,110; 4,517,288; and 4,837,168). For a review of the general immunoassays, see also *Methods in Cell Biology: Antibodies in Cell Biology*, volume 37 (Asai, ed. 1993); *Basic and Clinical Immunology* (Stites & Terr, eds., 7th ed. 1991). Immunological binding assays (or immunoassays) typically use an antibody that specifically binds to a protein or antigen of choice (in this case the ubiquitin ligation cascade protein or antigenic subsequence thereof). The antibody (e.g., anti- ubiquitin ligation cascade protein) may be produced by any of a number of means well known to those of skill in the art and as described above.

[0184] Immunoassays also often use a labeling agent to specifically bind to and label the complex formed by the antibody and antigen. The labeling agent may itself be one of the moieties comprising the antibody/antigen complex. Thus, the labeling agent may be a labeled ubiquitin ligation cascade protein or a labeled anti- ubiquitin ligation cascade protein antibody. Alternatively, the labeling agent may be a third moiety, such a secondary antibody, that specifically binds to the antibody/ antigen complex (a secondary antibody is typically specific to antibodies of the species from which the first antibody is derived). Other proteins capable of specifically binding immunoglobulin constant regions, such as protein A or protein G may also be used as the label agent. These proteins exhibit a strong non-immunogenic reactivity with immunoglobulin constant regions from a variety of species (*see, e.g.,* Kronval *et al., J. Immunol.* 111:1401-1406 (1973); Akerstrom *et al., J. Immunol.* 135:2589-2542 (1985)). The labeling agent can be modified with a detectable moiety, such as biotin, to which another molecule can specifically bind, such as streptavidin. A variety of detectable moieties are well known to those skilled in the art.

[0185] Throughout the assays, incubation and/or washing steps may be required after each combination of reagents. Incubation steps can vary from about 5 seconds to several hours, optionally from about 5 minutes to about 24 hours. However, the incubation time will depend upon the assay format, antigen, volume of solution, concentrations, and the like. Usually, the assays will be carried out at ambient temperature, although they can be conducted over a range of temperatures, such as 10°C to 40°C.

Non-competitive assay formats

[0186] Immunoassays for detecting ubiquitin ligation cascade protein in samples may be either competitive or noncompetitive. Noncompetitive immunoassays are assays in which the amount of antigen is directly measured. In one preferred "sandwich" assay, for example, the anti- ubiquitin ligation cascade protein antibodies can be bound directly to a solid substrate on which they are immobilized. These immobilized antibodies then capture the ubiquitin ligation cascade protein present in the test sample. Proteins thus immobilized are then bound by a labeling agent, such as a second ubiquitin ligation cascade protein antibody bearing a label. Alternatively, the second antibody may lack a label, but it may, in turn, be bound by a labeled third antibody specific to antibodies of the species from which the second antibody is derived. The second or third antibody is typically modified with a detectable moiety, such as biotin, to which another molecule specifically binds, e.g., streptavidin, to provide a detectable moiety.

Competitive assay formats

[0187] In competitive assays, the amount of ubiquitin ligation cascade protein present in the sample is measured indirectly by measuring the amount of a known, added (exogenous) ubiquitin ligation cascade protein displaced (competed away) from an anti- ubiquitin ligation cascade protein antibody by the unknown ubiquitin ligation cascade protein present in a sample. In one competitive assay, a known amount of ubiquitin ligation cascade protein is added to a sample and the sample is then contacted with an antibody that specifically binds to ubiquitin ligation cascade protein. The amount of exogenous ubiquitin ligation cascade protein bound to the antibody is inversely proportional to the concentration of ubiquitin ligation cascade protein present in the sample. In a particularly preferred embodiment, the antibody is immobilized on a solid substrate. The amount of ubiquitin ligation cascade protein bound to the antibody may be determined either by measuring the amount of ubiquitin ligation cascade protein present in ubiquitin ligation cascade protein/antibody complex, or alternatively by measuring the amount of remaining uncomplexed protein. The amount of ubiquitin ligation cascade protein may be detected by providing a labeled ubiquitin ligation cascade molecule.

[0188] A hapten inhibition assay is another preferred competitive assay. In this assay the known ubiquitin ligation cascade protein is immobilized on a solid substrate. A known amount of anti- ubiquitin ligation cascade protein antibody is added to the sample, and the sample is then contacted with the immobilized ubiquitin ligation cascade protein. The

amount of anti- ubiquitin ligation cascade protein antibody bound to the known immobilized ubiquitin ligation cascade protein is inversely proportional to the amount of ubiquitin ligation cascade protein present in the sample. Again, the amount of immobilized antibody may be detected by detecting either the immobilized fraction of antibody or the fraction of the antibody that remains in solution. Detection may be direct where the antibody is labeled or indirect by the subsequent addition of a labeled moiety that specifically binds to the antibody as described above.

Cross-reactivity determinations

[0189] Immunoassays in the competitive binding format can also be used for crossreactivity determinations. For example, a ubiquitin ligation cascade protein can be immobilized to a solid support. Proteins are added to the assay that compete for binding of the antisera to the immobilized antigen. The ability of the added proteins to compete for binding of the antisera to the immobilized protein is compared to the ability of the ubiquitin ligation cascade protein to compete with itself. The percent crossreactivity for the above proteins is calculated, using standard calculations. Those antisera with less than 10% crossreactivity with each of the added proteins listed above are selected and pooled. The cross-reacting antibodies are optionally removed from the pooled antisera by immunoabsorption with the added considered proteins, e.g., distantly related homologs.

[0190] The immunoabsorbed and pooled antisera are then used in a competitive binding immunoassay as described above to compare a second protein, thought to be perhaps an allele or polymorphic variant of a ubiquitin ligation cascade protein, to the immunogen protein. In order to make this comparison, the two proteins are each assayed at a wide range of concentrations and the amount of each protein required to inhibit 50% of the binding of the antisera to the immobilized protein is determined. If the amount of the second protein required to inhibit 50% of binding is less than 10 times the amount of the ubiquitin ligation cascade protein that is required to inhibit 50% of binding, then the second protein is said to specifically bind to the polyclonal antibodies generated to ubiquitin ligation cascade immunogen.

Other assay formats

[0191] Western blot (immunoblot) analysis is used to detect and quantify the presence of ubiquitin ligation cascade protein in the sample. The technique generally comprises separating sample proteins by gel electrophoresis on the basis of molecular weight,

transferring the separated proteins to a suitable solid support, (such as a nitrocellulose filter, a nylon filter, or derivatized nylon filter), and incubating the sample with the antibodies that specifically bind the ubiquitin ligation cascade protein. The anti- ubiquitin ligation cascade protein antibodies specifically bind to the ubiquitin ligation cascade protein on the solid support. These antibodies may be directly labeled or alternatively may be subsequently detected using labeled antibodies (e.g., labeled sheep anti-mouse antibodies) that specifically bind to the anti- ubiquitin ligation cascade protein antibodies.

[0192] Other assay formats include liposome immunoassays (LIA), which use liposomes designed to bind specific molecules (e.g., antibodies) and release encapsulated reagents or markers. The released chemicals are then detected according to standard techniques (*see* Monroe *et al.*, *Amer. Clin. Prod. Rev.* 5:34-41 (1986)).

Reduction of non-specific binding

[0193] One of skill in the art will appreciate that it is often desirable to minimize non-specific binding in immunoassays. Particularly, where the assay involves an antigen or antibody immobilized on a solid substrate it is desirable to minimize the amount of non-specific binding to the substrate. Means of reducing such non-specific binding are well known to those of skill in the art. Typically, this technique involves coating the substrate with a proteinaceous composition. In particular, protein compositions such as bovine serum albumin (BSA), nonfat powdered milk, and gelatin are widely used with powdered milk being most preferred.

Labels

[0194] The particular label or detectable group used in the assay is not a critical aspect of the invention, as long as it does not significantly interfere with the specific binding of the antibody used in the assay. The detectable group can be any material having a detectable physical or chemical property. Such detectable labels have been well-developed in the field of immunoassays and, in general, most any label useful in such methods can be applied to the present invention. Thus, a label is any composition detectable by spectroscopic, photochemical, biochemical, immunochemical, electrical, optical or chemical means. Useful labels in the present invention include magnetic beads (e.g., DYNABEADS^{TM}), fluorescent dyes (e.g., fluorescein isothiocyanate, Texas red, rhodamine, and the like), radiolabels (e.g., ^{3}H, ^{125}I, ^{35}S, ^{14}C, or ^{32}P), enzymes (e.g., horse radish peroxidase, alkaline phosphatase and

others commonly used in an ELISA), and colorimetric labels such as colloidal gold or colored glass or plastic beads (e.g., polystyrene, polypropylene, latex, etc.).

[0195] The label may be coupled directly or indirectly to the desired component of the assay according to methods well known in the art. As indicated above, a wide variety of labels may be used, with the choice of label depending on sensitivity required, ease of conjugation with the compound, stability requirements, available instrumentation, and disposal provisions.

[0196] Non-radioactive labels are often attached by indirect means. Generally, a ligand molecule (e.g., biotin) is covalently bound to the molecule. The ligand then binds to another molecules (e.g., streptavidin) molecule, which is either inherently detectable or covalently bound to a signal system, such as a detectable enzyme, a fluorescent compound, or a chemiluminescent compound. The ligands and their targets can be used in any suitable combination with antibodies that recognize ubiquitin ligation cascade protein, or secondary antibodies that recognize anti- ubiquitin ligation cascade protein.

[0197] The molecules can also be conjugated directly to signal generating compounds, e.g., by conjugation with an enzyme or fluorophore. Enzymes of interest as labels will primarily be hydrolases, particularly phosphatases, esterases and glycosidases, or oxidotases, particularly peroxidases. Fluorescent compounds include fluorescein and its derivatives, rhodamine and its derivatives, dansyl, umbelliferone, etc. Chemiluminescent compounds include luciferin, and 2,3-dihydrophthalazinediones, e.g., luminol. For a review of various labeling or signal producing systems that may be used, see U.S. Patent No. 4,391,904.

[0198] Means of detecting labels are well known to those of skill in the art. Thus, for example, where the label is a radioactive label, means for detection include a scintillation counter or photographic film as in autoradiography. Where the label is a fluorescent label, it may be detected by exciting the fluorochrome with the appropriate wavelength of light and detecting the resulting fluorescence. The fluorescence may be detected visually, by the use of electronic detectors such as charge coupled devices (CCDs) or photomultipliers and the like. Similarly, enzymatic labels may be detected by providing the appropriate substrates for the enzyme and detecting the resulting reaction product. Colorimetric or chemiluminescent labels may be detected simply by observing the color associated with the label. Thus, in various dipstick assays, conjugated gold often appears pink, while various conjugated beads appear the color of the bead.

[0199] Some assay formats do not require the use of labeled components. For instance, agglutination assays can be used to detect the presence of the target antibodies. In this case, antigen-coated particles are agglutinated by samples comprising the target antibodies. In this format, none of the components need be labeled and the presence of the target antibody is detected by simple visual inspection.

EXAMPLES

[0200] The following examples are offered to illustrate, but not to limit the claimed invention.

Example 1: separation of ubiquitinated proteins from non-ubiquitinated proteins.

[0201] Ubiquitinated proteins and non-ubiquitinated proteins from cell extracts were separated using three different methods.

Method 1: In vivo separation using S5a affinity chromatography

[0202] The S5a subunit of the 26S proteosome binds polyubiquitinated proteins. (Young *et al., J. Biol. Chem.* 273: 5461-5467 (1998)). Briefly, asynchronously growing HeLa or A549 cells were treated with the proteosome inhibitor MG132 and polyubiquitinated proteins were isolated using GST-S5a protein immobilized on agarose beads.

[0203] The full length human S5a cDNA was cloned in to the pGEX-4T-1 N-terminal glutathione-S-transferase (GST) fusion prokaryotic expression vector by standard methods, and the sequence integrity was confirmed by automated DNA sequencing. (Layfield *et al., Proteomics* 1:773-777 (2001)). Prior to cell harvesting, asynchronously growing cells were treated with proteosome inhibitor, MG132 (10 μM) for 4 h to allow accumulation of ubiquitinated proteins. Polyubiquitinated cellular proteins were extracted using a GST-S5a fusion attached to agarose beads after lysing the cells in the presence of TBS buffer containing 1% triton X-100, iodoacetamide (10 mM) and a cocktail of protease inhibitors. Extracted proteins were visualized either by silver staining or by immunoblotting of 1D or 2D gels. Affinity extraction using GST alone matrix served as a control.

[0204] GST-S5a matrix binds cellular proteins with at least 4 ubiquitin chains and also proteins that contain intrinsic Ub-like domains, *e.g.*, hHR23B. (Layfield *et al. Proteomics* 1:773-777 (2001)). Results are shown in Fig. 2. Total cell lysate, protein eluted from GST control beads, and protein eluted from GST-S5a beads were analyzed by SDS-PAGE followed by Western blotting. Ubiquitinated proteins were detected by probing with antibodies specific for ubiquitin.

Method 2: In vivo separation using FLAG epitope affinity chromatography

[0205] Stable HeLa and A549 cell lines expressing flag tagged ubiquitin were created using standard retroviral vectors. (*See, e.g.*, Lorens *et al., Mol. Ther.* 1:438-447 (2000)). Briefly, FLAG-tagged polyubiquitinated proteins were accumulated by treating the cells using MG132 (proteosome inhibitor) and isolated using anti-FLAG agarose beads.

[0206] Prior to cell harvesting, asynchronously growing HeLa and A549 cells were treated with proteosome inhibitor, MG132 (10 μM) for 4 h to allow accumulation of ubiquitinated proteins. Polyubiquitinated cellular proteins were extracted using a anti-FLAG agarose beads after lysing the cells in the presence of TBS buffer containing 1% triton X-100, iodoacetamide (10 mM) and a cocktail of protease inhibitors. Extracted proteins were visualized either by silver staining or by immunoblotting of 1D or 2D gels. Expression of GFP alone served as control.

[0207] Ubiquitinated proteins were detected with an antibody specific for the FLAG epitope used to label ubiquitin. As a control, lysates from cells infected with a control virus, *i.e.*, GFP, were also analyzed.

Method 3: In vitro separation using Ni-NTA affinity chromatography

[0208] Briefly, *in vitro* ubiquitin conjugation reactions were done using HeLa cell lysates in the presence of $(his)_6$-tagged UB and ATP. His-Ubiquitin labeled proteins were isolated by Ni-NTA affinity chromatography (IMAC) and subjected for MS/MS analysis after digesting the whole affinity extract using trypsin.

[0209] A HeLa cell line established from an adenocarcinoma cervix was cultured in modified DMEM medium, supplemented with 10% FBS plus 1% P/S in a 37°C incubator with 5% CO_2. The adherent asynchronously growing HeLa cells were harvested after 48 h by trypsinization followed by low speed centrifugation (1200 rpm, 10 min). Cells were washed twice with PBS and then lysed in HEPES buffer (50mM, pH 7.5) containing KCl (100 mM), 0.2% NP40, TCEP (0.5 mM) and $MgCl_2$ (5 mM) in a nitrogen bomb. After centrifugation (3000 rpm, 10 min), the supernatant was subjected to high-speed centrifugation at 105,000 X g for one hour at 4°C (HeLa S-100 fraction).

[0210] For the *in vitro* conjugation reaction, 2.3 ml HeLa S-100 fraction (5 mg/ml protein content) was used along with purified His-Ub (1mg); a deubiquitinating enzyme inhibitor, UB-CHO (5 μM); and a proteosome inhibitor, MG132 (10 μM). The reaction was started by adding an ATP regeneration system and continued for 1 h at room temperature. A reaction

mixture without an ATP regeneration system and cell lysates pre-treated with iodoacetamide (20 mM) to block endogenous E1/E2 activity served as controls.

[0211] After the reaction, his-tagged polyubiquitinated proteins were isolated by conventional IMAC methods using Ni-NTA resin (Sigma). All affinity extractions were done in qudruplicate and the washed beads containing protein complexes were stored at 4°C until analyzed. For 1D-gel and Western analysis, washed beads were separated by SDS-PAGE on 4-20% Tris glycine PAGE system and blotted as per the manufacturer's recommendations (Novex, SanDiego, CA). Proteins electroblotted onto PVDF membranes were probed with anti-his (Santa Cruz Biotechnology) or anti-Ub (Zymed) antibodies and developed using an ECL plus enhanced chemiluminescence reagent kit (Amersham Pharmacia) followed by detection on a ECL hyperfilm.

[0212] Trypsinization of whole affinity extract was accomplished by carrying out on-bead digestion using 100 μL of 8 M urea (Sigma Chemical Co., St. Louis, MO) containing methylamine (20 mM), LiCl (1 M), EDTA (2 mM) in 0.1 M tris buffer (pH 8.5). Proteins on the beads were reduced by DTT and carboxyamido-methylated by iodoacetamide. Lys-C endoproteinase (Boehringer Mannheim, Germany) was added into the protein solution (diluted to 4 M urea) to a final concentration of 5 μg/ml, and incubated at 37°C for 15 hrs. Then modified trypsin (Boehringer Mannheim, Germany) was added into the solution (diluted to 2 M urea) to a final concentration of 25 μg/ml and incubated at 37°C, overnight. The reaction was quenched by adding 1μl of glacial acetic acid.

[0213] The efficiency of *in vitro* ubiquitination using HeLa cell lysates in the presence of $(his)_6$-tagged Ub/ATP was found to be about 10 to 12% based on the amount of unreacted His-ubiquitin left after 1 h reaction (determined by gel densitometry).

[0214] Figures 2 and 3 show 1D gel analysis of the affinity extracts derived after trapping His-Ub tagged cellular proteins by IMAC. Ubiquitination of proteins was confirmed by immunoblot using anti-his antibodies and anti-Ubiquitin antibodies.

Example 2: isolation of ubiquitinated proteins from each other.

[0215] For micro-capillary LC MS/MS analysis, 5 μL of the trypsin digest, was desalted using a C18 guard column (Vydac, Hesperia, CA) offline, and then injected directly onto an in-house prepared biphasic microcapillary hplc column (16 cm long, 75 μm i.d. fused silica PicoFrit with 15 μm tip, New Objective Inc., Woburn MA) attached to an Ultimate microcapillary HPLC (LC packings, San Francisco, CA). The first phase of the column contains 12 cm of reverse-phase C18 media (Aquasil $C1_{18}$, 3 um i.d., Thermo Hypersil-

Keystone Scientific, Bellefonte, PA) and the second phase contains 4 cm of strong-cation exchange (SCX) media (Polysulfoethyl Aspartamide, PolyLC, Columbia, MA). Separation of tryptic peptide mixtures were accomplished using a step gradient of SCX buffer system (5 mM, 10 mM, 20 mM, 50 mM, 100 mM and 500 mM NH_4OAc containing 5% ACN and 0.5% AcOH, pH 4) in one dimension before running a reversed phase buffer system (A = 5% ACN + 0.5% AcOH; B = 80% ACN + 0.5% AcOH) in the second dimension. Each run was carried out for 2.5 h with a flow rate of ca. 200 nl/min. Samples emerging out from the column were detected using Q-TOF tandem mass spectrometer (Micromass, Manchester, UK).

Example 3: identification of isolated, ubiquitinated proteins.

[0216] Protocols for isolation and identification of ubiquitinated proteins are depicted in Figs. 4 and 5. Before conducting the database search using the collected MS/MS data, a CGI Perl program called "MS2Filter" was used to pre-filter the raw data. This program digitally compares all experimental MS/MS spectra in quadruplicate runs with those found in each of the quadruplicate control runs and eliminates them before initiating the database search. This pre-filtering of the MS/MS data significantly reduced the bulk of the MS/MS data and increased the speed of protein identification. The remaining tandem mass spectra after pre filtering were analyzed using TurboSequest software (Thermofinnigan, San Jose, CA), which uses the SEQUEST algorithm, against a human protein sequence database derived from the NCBI non redundant human database (INFORMATICS Websites used: ncbi.nlm.nih.gov:80/entrez; isrec.isbsib.ch/software/PFSCAN_form.html; blocks.fhcrc.org/blocks/blocks_search.html). SEQUEST results were summarized and further analyzed using a separate database called "MEDUSA" (Gururaja *et al., J. Proteome Res.* 1:253-61 (2002)). MEDUSA is a proprietary web-based Oracle 8.0 database that we have designed and coded that stores and evaluates mass spectrometry data and SEQUEST output. It runs independently of SEQUEST.

[0217] After acquiring MS/MS data followed by database search using SEQUEST software and data mining with the aid of MS2Filter and MEDUSA as outlined in Fig. 5, approximately 240 unique proteins arising out of experimental affinity extract were identified. All affinity extractions and MS runs were carried out in quadruplicate and only proteins present in at least 2x were selected.

[0218] A total of 22 proteins belonging to 26S proteosome subunit (Table 1) and about 20 different proteins of ubiquitination system such as various E2 and E3 enzymes,

deubiquitinating enzymes and ubiquitin itself (Table 2) were identified. Interestingly, three proteins which have ubiquitin like domains such as RAD23, BAT3, GDX proteins (Table 2) were identified.

TABLE 1

PROTEIN ID BY 2D-LC/MS/MS: ***Proteins Belonging To Proteosome Subunits***

Proteins	Peptides	Reproducibility
26S proteasome regulatory subunit, S4, ATPase 1	13	4 out of 4
TBP1, Tat binding protein 1, 26S proteasome subunit, ATPase	13	4 out of 4
26S proteasome regulatory complex chain p44.5, non ATPase	9	4 out of 4
26S proteasome regulatory subunit 7 (MSS1 protein)	8	4 out of 4
26S proteasome regulator subunit 8 (TRIP1 protein)	8	4 out of 4
26S proteasome subunit chain 5a	7	4 out of 4
26S proteasome regulatory subunit 3, non-ATPase	6	4 out of 4
26S proteasome regulatory subunit S12, non-ATPase	5	4 out of 4
26S proteasome regulatory subunit, non-ATPase 1	5	4 out of 4
26S proteasome associated pad1 homolog	4	3 out of 4
26S proteasome related 55.1 protein homolog	3	4 out of 4
Proteasome regulatory particle subunit p44S10	3	3 out of 4
Proteasome 26S subunit, ATPase 5 (SUG1)	3	4 out of 4
Proteasome 26S subunit, non ATPase, 3	3	3 out of 4
Proteasome (subunit 7, alpha type, isoform 4	3	3 out of 4
Proteasome 26S subunit, ATPase homolog, 4	3	3 out of 4
p67 similar to proteasome 26S subunit, non-ATPase, 2	2	4 out of 4
Proteasome 26S subunit, non-ATPase, 7 (Mov34 homolog)	2	3 out of 4
Proteasome subunit, alpha type 6	2	3 out of 4
Proteasome 26S subunit, non-ATPase, 5	1	2 out of 4
Proteasome subunit, beta type, 1	1	2 out of 4
Proteasome subunit, beta type, 7	1	2 out of 4

TABLE 2

PROTEIN ID BY 2D-LC/MS/MS: *Ubiquitination System Proteins*

Proteins	Peptides	Reproducibility
KIAA0312 (contains HECT domain and E3 ligase activity)	20	4 out of 4
CGI-70 protein (Ubiquitin carboxyl terminal hydrolase family, 1)	14	4 out of 4
KIAA1578 (similar to ubiquitin-protein ligase 1, putative)	9	3 out of 4
Hyd protein (contains HECT domain and E3 ligase activity)	8	4 out of 4
Ubiquitin-conjugating BIR domain enzyme APOLLON	6	3 out of 4
Ring finger protein 20	4	3 out of 4
Ubiquitin conjugating enzyme	3	4 out of 4
FLJ12878, likely ortholog of mouse ubiquitin conjugating enzyme E2-230K	3	4 out of 4
KIAA1734, likely ortholog of mouse ubiquitin-conjugating enzyme E2 230K	3	4 out of 4
Ubiquitin C-terminal hydrolase, UBCH37	2	3 out of 4
HSPC150 protein similar to ubiquitin conjugating enzyme	2	3 out of 4
UBCF_Human ubiquitin conjugating enzyme E2-24KDa	2	2 out of 4
Ubiquitin	1	3 out of 4
Progestin induced protein; ubiquitin-protein ligase	1	2 out of 4
Ubiquitin conjugating enzyme E2N (similar to yeast UBC13)	1	2 out of 4
Ubiquitin specific protease 11	1	2 out of 4
p45 SKP2 like F-box protein	1	2 out of 4
Autocrine motility factor receptor, gp78, E3 ligase for ER	1	2 out of 4
UBIQUITIN LIKE / UBIQUITIN INTERACTING PROTEINS		
BAT3, human ubiquitin like protein	5	4 out of 4
UBL4, human ubiquitin like protein GDX	1	2 out of 4
HHR23B, RAD23 human homolog B ubiquitin like protein	1	2 out of 4
Josephin MJD1, similar to ataxin-3 interacts with Ub-like domain of HHR23B	3	3 out of 4

[0219] Nine different known ubiquitin protein substrates such as histone H2A, IRS-2, P13 kinase, CENP-E etc. (Table 3). Proteins not previously known to be ubiquitin substrates (Table 4) were also identified.

TABLE 3.

PROTEIN ID BY 2D-LC/MS/MS: *Known Ubiquitinated Proteins*

Proteins	Peptides	Reproducibility
EPS15, epidermal growth factor receptor pathway substrate 15	3	4 out of 4
hnRNP U, heterogeneous ribonuclear particle protein U	2	4 out of 4
IRS-2, insulin receptor substrate 2	2	4 out of 4
H2A histone family member	5	3 out of 4
PI3 kinase, phosphoinositide-3-kinase, class 2 protein	1	3 out of 4
p120 catenin isoform 413	2	3 out of 4
CENP-E protein, centromere protein E	2	2 out of 4
Ras GTPase-activating protein IQGAP1	2	2 out of 4
Ras homolog gene family, member A	2	2 out of 4

TABLE 4

PROTEIN ID BY 2D-LC/MS/MS: *Unknown Proteins for Ubiquitination*

Proteins	Peptides	Reproducibility
Rb associated factor 600	9	4 out of 4
Oncogene EMS1, human SRC substrate CORTACTIN	6	4 out of 4
Glucose regulated protein 58	4	4 out of 4
Protein disulfide isomerase	4	4 out of 4
Human thioredoxin related protein	4	4 out of 4
Valosin containing protein	2	4 out of 4
OTU like cysteine protease	3	4 out of 4
Splicing factor 313 subunit	3	4 out of 4
Activating signal cointegrator 1	2	4 out of 4
Receptor interacting serine threonine kinase 2 (RIP2 kinase)	2	3 out of 4
Multifunctional protein CAD	2	3 out of 4
ATP binding cassette protein C11	2	3 out of 4
LAR interacting protein LIP1b	2	3 out of 4
DAF, decay accelerating factor	1	3 out of 4
G protein-coupled receptor kinase interactor - 1	1	3 out of 4
BRCT domain containing protein	1	3 out of 4
Reticulocalbin 2, ER calcium binding protein	1	3 out of 4
Opioid growth factor receptor	1	3 out of 4
KARP-1 binding protein	1	2 out of 4
Rab 11 binding protein	1	2 out of 4
C-myc binding protein	1	2 out of 4
Cell cycle division protein, CDC10	1	2 out of 4
elF 2B, transcription initiation factor 2B	1	2 out of 4
TBX 22, T-box transcription factor 22	1	2 out of 4
PAK interacting factor beta	1	2 out of 4
Elongation factor B	1	2 out of 4

[0220] Once made a ubiquitin map, derived from in vitro or in vivo ubiquitination, has several uses. In one embodiment a ubiquitin map is used differentially, to detect changes, with and without prior exposure of the cell (or sub cellular) lysate or the live cell, to other agents, such as small molecules, drugs, dominant negative forms of proteins, antisense to mRNA coding proteins of interest, Si RNA knocking out RNA for proteins of interest, overexpression of proteins of interest.

[0221] In addition the ubiquitin map is used as a readout for the discovery of changes in the levels of individual ubiquitinated proteins upon some prior treatment of the cell, including synchronizing cells at individual stages of the cell cycle. The map can be used to compare ubiquitin maps between different cell types (tumor vs. normal, for example).

[0222] In one embodiment maps can be combined to look for the effect of forcing one modification (such as ubiquitination or sumoylation) on another map (neddylation for

example) in any combination. The map can be used as a readout for a cellular or in vitro screen of effectors of ubiquitination (or sumoylation, neddylation etc.).

[0223] It is understood that the examples and embodiments described herein are for illustrative purposes only and that various modifications or changes in light thereof will be suggested to persons skilled in the art and are to be included within the spirit and purview of this application and scope of the appended claims. All publications, patents, and patent applications cited herein are hereby incorporated by reference in their entirety for all purposes.

WHAT IS CLAIMED IS:

1. A method for the identification of ubiquitinated proteins comprising:

a. separating ubiquitinated proteins from non-ubiquitinated proteins;

b. isolating said ubiquitinated proteins from each other; and

c. identifying at least a plurality of said ubiquitinated proteins.

2. A method according to claim 1, further comprising generating a ubiquitin map from said identified proteins.

3. A method according to claim 1, further comprising providing a first population of cells and lysing said cells to provide a pool of proteins for said separation.

4. A method according to claim 3, further comprising:

a. providing a second population of cells and lysing said cells to provide a second pool of proteins for said separation.;

b. separating ubiquitinated proteins from non-ubiquitinated proteins;

c. isolating said ubiquitinated proteins from each other;

d. identifying at least a plurality of said ubiquitinated proteins; and

e. comparing the identified proteins from said first population with the identified population of said second population.

5. A method according to claim 4, wherein said first population comprises cancer cells and said second population comprises normal cells.

6. A method according to claim 4, wherein said first and second population are from the same cell type, and said first population has been treated with a drug.

7. A method according to claim 1, wherein said isolation is done by 2D gel analysis.

8. A method according to claim 1, wherein said isolation is done by purifying said proteins.

9. A method according to claim 1, wherein said separating comprises:

a. labeling said ubiquitinated proteins with a label selective for a ubiquitin moiety; and

b. purifying said ubiquitinated proteins by binding said label to a complementary binding partner.

10. The method according to claim 9, wherein said labeling occurs in vitro.

11. The method according to claim 10, wherein said labeling comprises:

a. preparing a reaction mixture comprising:

i) a cellular extract or subcellular fraction; and

ii) labeled ubiquitin moiety, whereby substrates in said cellular extract are ubiquitinated forming labeled ubiquitinated molecules;

b. contacting said labeled ubiquitinated molecules with a binding partner that selectively binds said label, thereby removing said labeled ubiquitinated proteins.

12. The method according to claim 11, wherein said label is selected from the group consisting of a His-tag, Glutathione-S-transferase, Flag-tag, HA-tag, affinity tag, epitope tag and biotin.

13. The method according to claim 11, further comprising providing proteosome inhibitors in said reaction mixture.

14. The method according to claim 13, wherein said proteosome inhibitors are selected from the group consisting of MG132, PS341 and epoxymicin.

15. The method according to claim 11, further comprising providing deubiquitinase inhibitors.

16. The method according to 15, wherein said deubiquitinase inhibitor is ubiquitin aldehyde.

17. The method according to claim 13, further comprising providing an ATP regenerating system in said reaction mixture.

18. The method according to claim 9, wherein said labeling occurs in vivo.

19. The method according to claim 18, wherein said labeling comprises providing cells with a labeled ubiquitin moiety whereby ubiquitin targets are labeled in vivo with said labeled ubiquitin moiety.

20. The method according to claim 19, wherein said labeled moiety is selected from a His-tag ubiquitin, Glutathione-S-transferase ubiquitin and Flag-tag ubiquitin HA-tagged ubiquitin, affinity tag ubiquitin, epitope tag ubiquitin and biotin tagged ubiquitin.

21. The method according to claim 20, further comprising incubating cells with proteosome inhibitors.

22. The method according to claim 21, wherein said proteosome inhibitors are selected from the group consisting of MG132, PS341 and epoxymicin.

23. The method according to claim 20, further comprising incubating cells with deubiquitinase inhibitors.

24. The method according to claim 23, wherein said deubiquitinase inhibitor is ubiquitin aldehyde.

25. A method according to claim 1, wherein said separating comprises:

a. performing a first dimensional separation of said ubiquitin proteins whereby said first dimensional separation is affinity purification; and

b. performing a second dimensional separation of said ubiquitin proteins whereby said second dimensional separation is LC-MS/MS.

26. The method according to claim 19, further comprising obtaining the sequence of said ubiquitinated proteins.

27. The method according to claim 25, wherein prior to b), said method comprises cleaving said affinity purified ubiquitinated proteins.

28. The method according to claim 27, wherein said cleaving is by a protease.

29. The method according to claim 28, wherein said protease is selected from the group consisting of lys C endo protease and trypsin.

30. A method for diagnosing a disorder associated with ubiquitination comprising:

a. obtaining a ubiquitin map from cells from an individual suspected of having a disorder associated with ubiquitin;

b. comparing said ubiquitin map from said individual with a ubiquitin map from a control, whereby a difference is indicative of a disorder associated with ubiquitination.

31. A method of identifying a disorder associated with ubiquitin, the method comprising

a. obtaining a ubiquitin map from cells from an individual with a disorder suspected to be associated with ubiquitin;

b. comparing said ubiquitin map from said individual with a ubiquitin map from a control, whereby a difference is indicative of a disorder associated with ubiquitination.

32. The method of claim 31, wherein ubiquitin maps are obtained from a plurality of individuals with the disorder suspected to be associated with ubiquitin.

33. A method of identifying a target protein of a ubiquitin pathway protein, wherein the identification takes place in the presence of a modulator of the ubiquitin pathway protein, the method comprising

a. separating ubiquitinated proteins from non-ubiquitinated proteins;

b. isolating said ubiquitinated proteins from each other; and

c. identifying at least a plurality of said ubiquitinated proteins;

wherein a difference in ubiquitination in the absence of the modulator identifies the protein as a target for a ubiquitin pathway protein.

34. The method of claim 33, wherein the ubiquitin pathway protein is ubiquitin ligation cascade component.

35. The method of claim 34, wherein the modulator is an inhibitor of the ubiquitin ligation cascade component.

36. The method of claim 35, wherein the inhibitor is selected from the group consisting of a dominant negative mutant, small molecules, siRNA, and antisense RNA.

37. The method of claim 34, wherein the modulator is an activator of the ubiquitin ligation cascade component.

38. The method of claim 37, wherein the activator is an overexpressed ubiquitin ligation cascade component.

39. The method of claim 33, wherein the ubiquitin pathway protein is deubiquitinating protein.

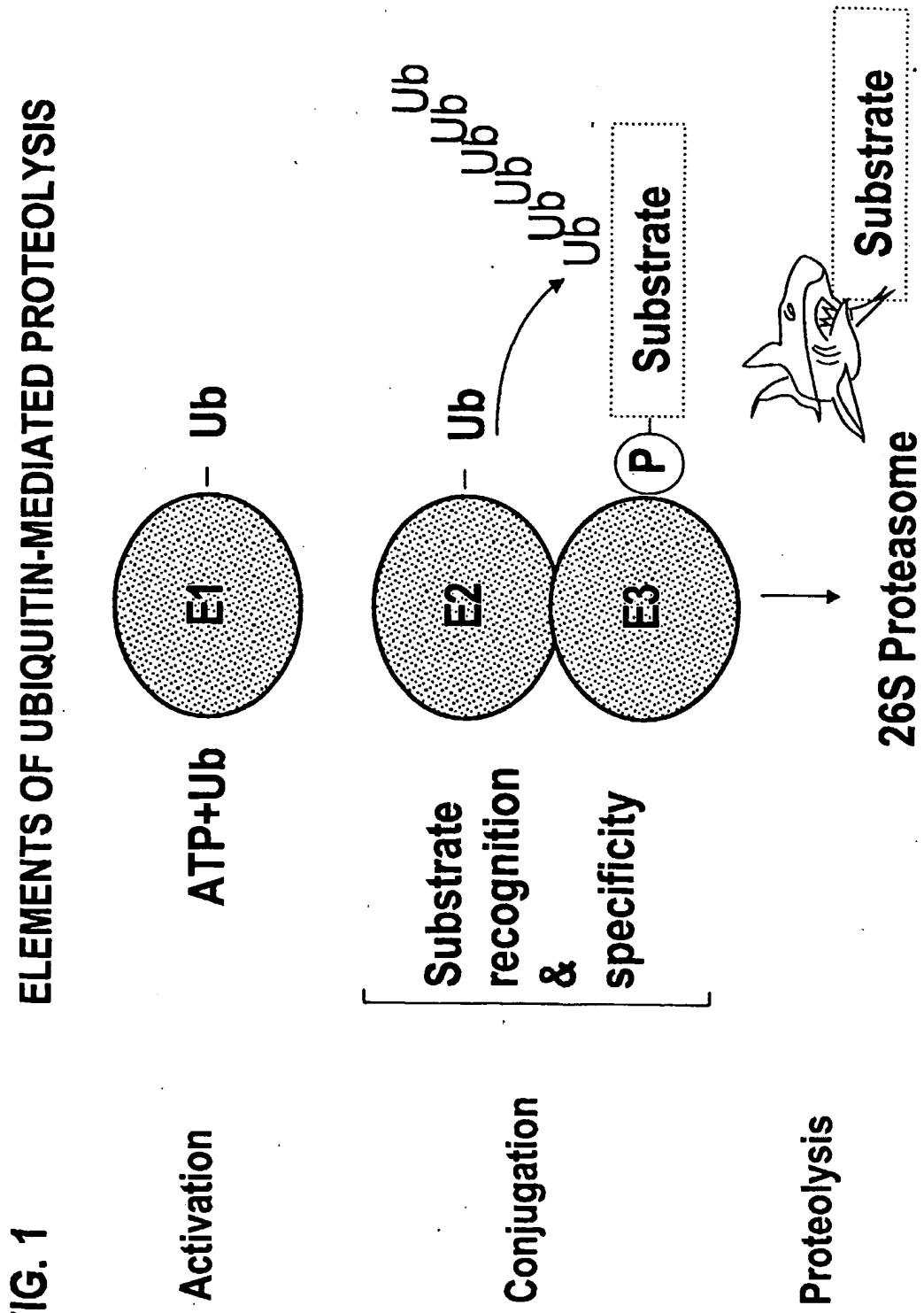
40. The method of claim 39, wherein the modulator is an inhibitor of the deubiquitinating protein.

41. The method of claim 40, wherein the inhibitor is selected from the group consisting of a dominant negative mutant, small molecules, siRNA, and antisense RNA.

42. The method of claim 39, wherein the modulator is an activator of the deubiquitinating protein.

43. The method of claim 42, wherein the activator is an overexpressed deubiquitinating protein.

FIG. 1



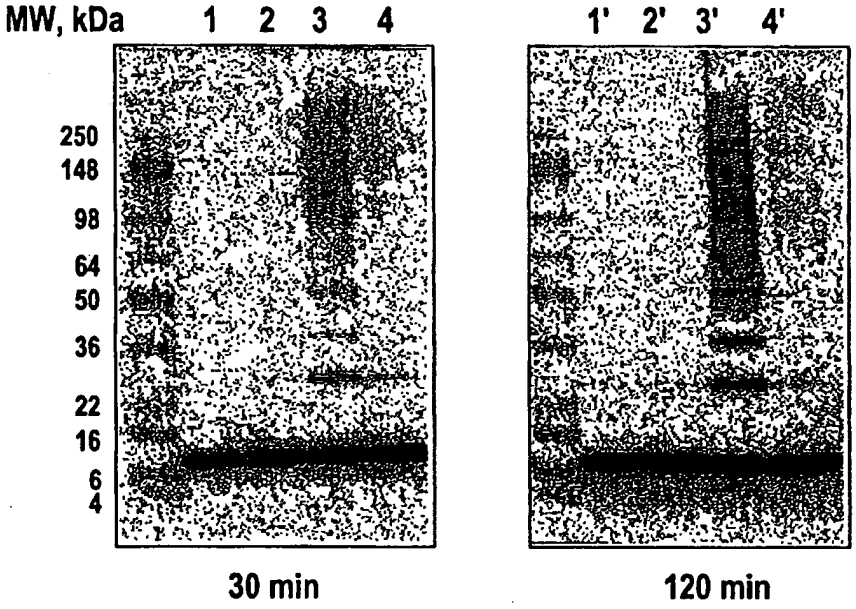


FIG. 2

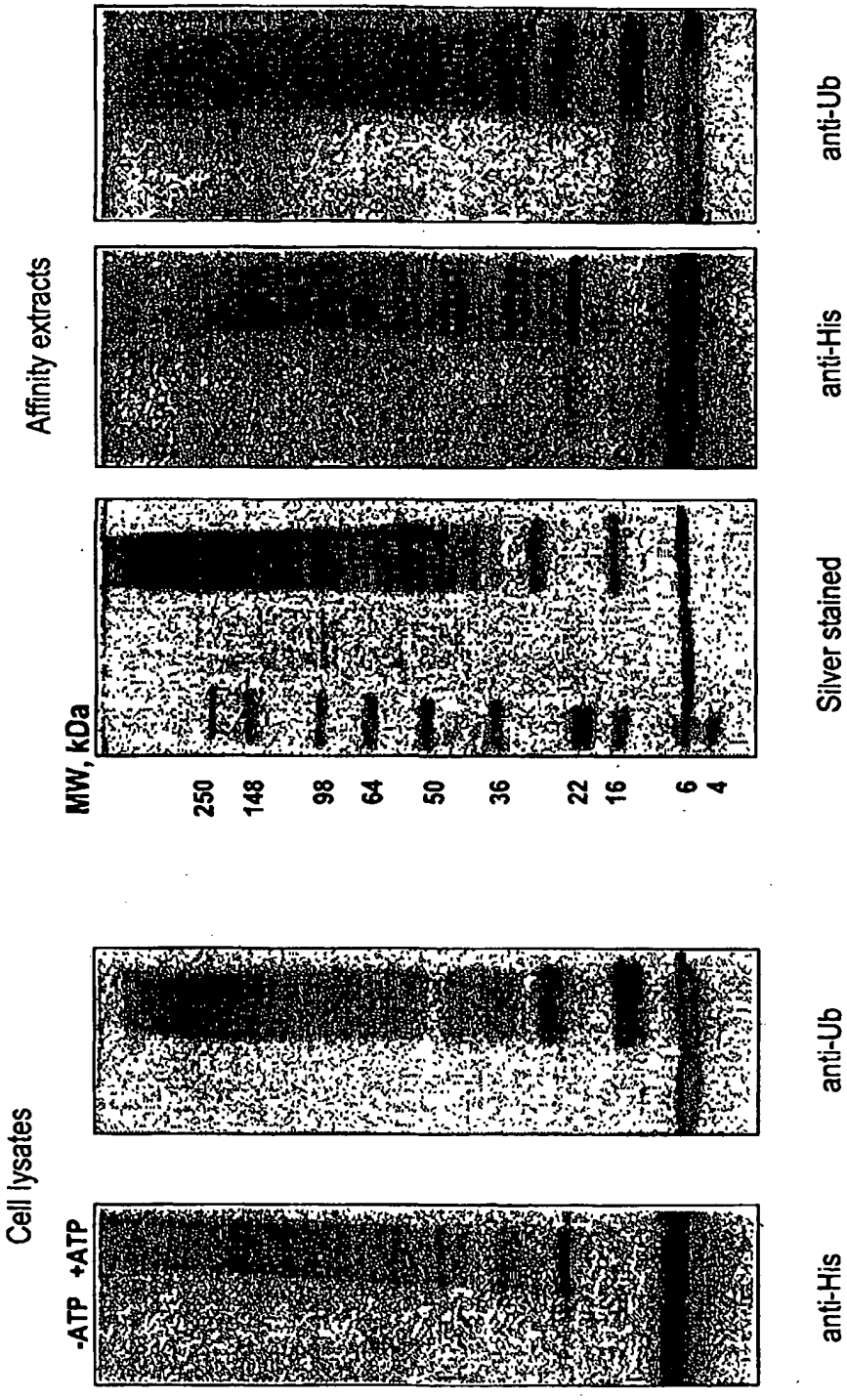

Affinity extracts
MW, kDa
250
148
98
64
50
36
22
16
6
4
Silver stained
anti-His
anti-Ub
Cell lysates
-ATP +ATP
anti-His
anti-Ub

FIG. 3

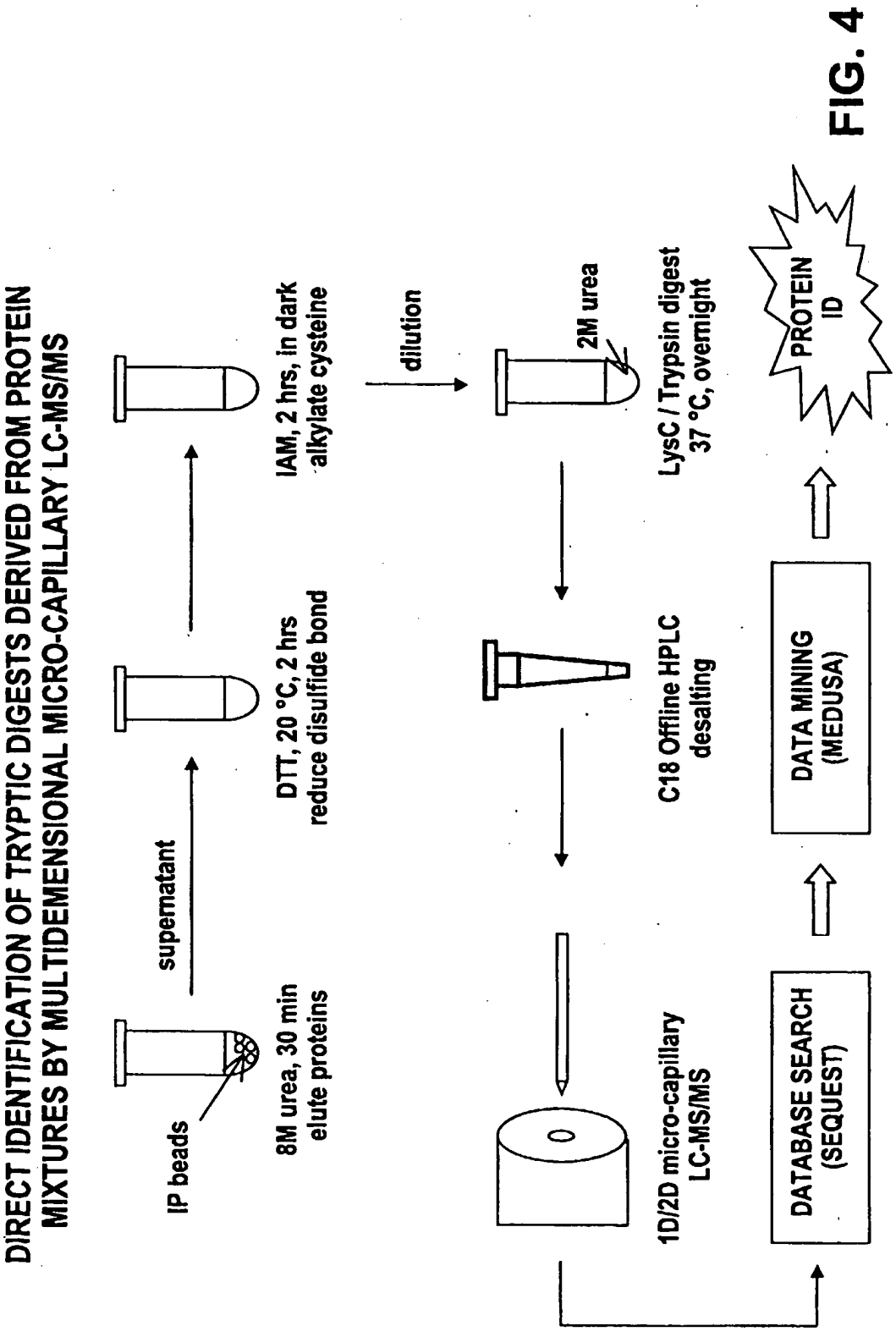

DIRECT IDENTIFICATION OF TRYPTIC DIGESTS DERIVED FROM PROTEIN MIXTURES BY MULTIDEMENSIONAL MICRO-CAPILLARY LC-MS/MS
IP beads
8M urea, 30 min elute proteins
supernatant
DTT, 20 °C, 2 hrs reduce disulfide bond
IAM, 2 hrs, in dark alkylate cysteine
dilution
2M urea
LysC / Trypsin digest 37 °C, overnight
C18 Offline HPLC desalting
1D/2D micro-capillary LC-MS/MS
DATABASE SEARCH (SEQUEST)
DATA MINING (MEDUSA)
PROTEIN ID

FIG. 4

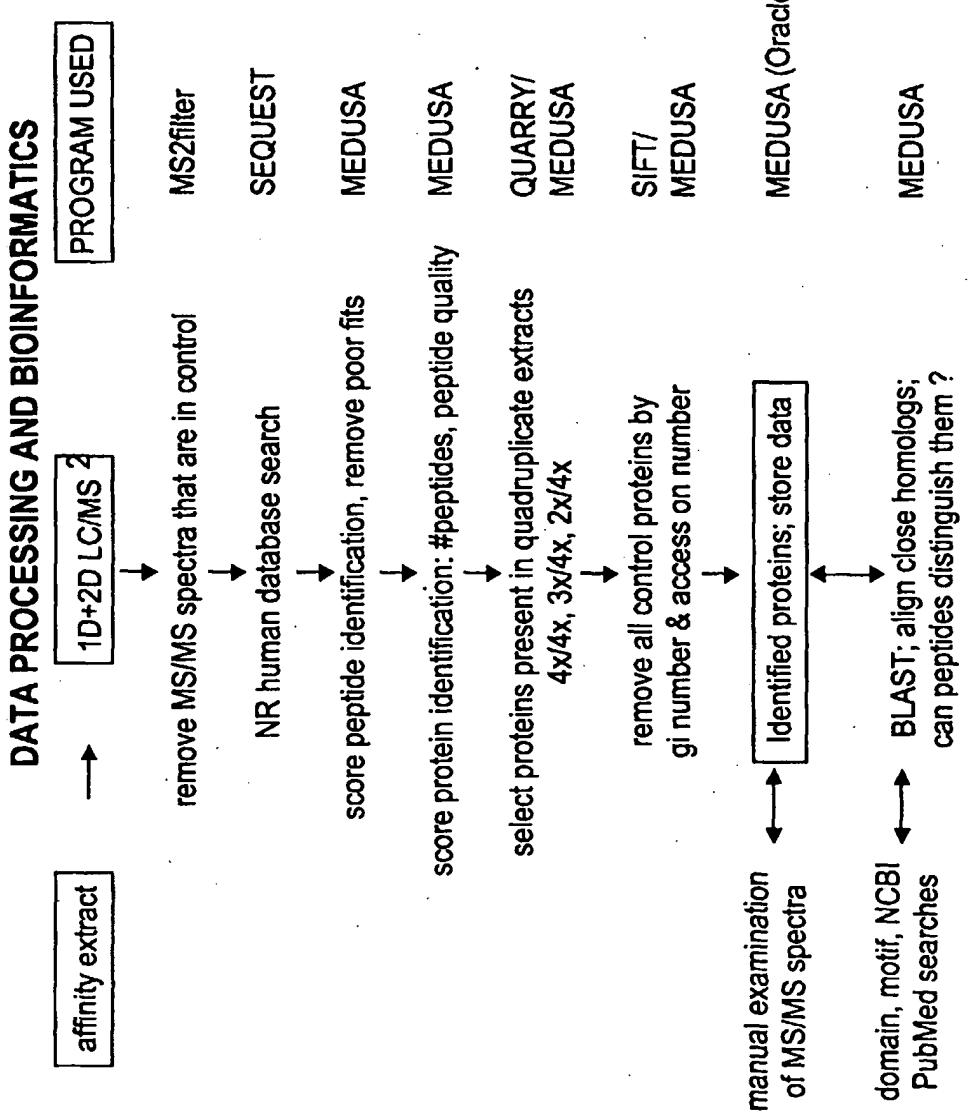


FIG. 5

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US03/17503

A. CLASSIFICATION OF SUBJECT MATTER
IPC(7) : C12N 9/00, 15/52, 15/11; C07H 21/04
US CL : 435/183, 243, 254.2, 320, 325, 410, 455; 536/23.1, 23.2, 24.3, 24.31, 24.33.
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 435/183, 243, 254.2, 320, 325, 410, 455; 536/23.1, 23.2, 24.3, 24.31, 24.33.

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
NONE

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EAST, PALM

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 5,976,849 A (HUSTAD et al.) 02 November 1999 (02.11.1999), col. 3, line 25-col. 4, line 24	1-43
A	US 5,952,481 A (MARKHAM et al.) 14 September 1999 (14.09.1999), col. 1, line 1-col. 14, line 45.	1-43
A	US 5,968,761 A (ROLFE et al.) 19 October 1999 (19.10.1999), col. 2, line 26-col. 8, line 60.	1-43

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
22 August 2003 (22.08.2003)	29 AUG 2003
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US Commissioner for Patents P.O. Box 1450 Alexandria, Virginia 22313-1450 Facsimile No. (703)305-3230	Authorized officer Phuong T. Do Telephone No. 703-308-0196

Form PCT/ISA/210 (second sheet) (July 1998)